

GenCode version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compend Ltd.

OW protein - protein search, using sw model

Run on: March 26, 2003, 16:44:40 ; Search time 9.59302 Seconds

(Without alignments)
46,007 Million cell updates/sec

Title: US-10-010-667a-2_COPY_14_28

Sequence: 86 WKPKPRNLEDDYL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 49 summaries

Database : Issued Patents AA.*
1: /cgn2_6/prodata/1/aa/5A.COMB.pep.*
2: /cgn2_6/prodata/1/aa/5B.COMB.pep.*
3: /cgn2_6/prodata/1/aa/6A.COMB.pep.*
4: /cgn2_6/prodata/1/aa/6B.COMB.pep.*
5: /cgn2_6/prodata/1/aa/6C.COMB.pep.*
6: /cgn2_6/prodata/1/aa/6D.COMB.pep.*
7: /cgn2_6/prodata/1/aa/6E.COMB.pep.*
8: /cgn2_6/prodata/1/aa/6F.COMB.pep.*
9: /cgn2_6/prodata/1/aa/6G.COMB.pep.*
10: /cgn2_6/prodata/1/aa/6H.COMB.pep.*
11: /cgn2_6/prodata/1/aa/6I.COMB.pep.*
12: /cgn2_6/prodata/1/aa/6J.COMB.pep.*
13: /cgn2_6/prodata/1/aa/6K.COMB.pep.*
14: /cgn2_6/prodata/1/aa/6L.COMB.pep.*
15: /cgn2_6/prodata/1/aa/6M.COMB.pep.*
16: /cgn2_6/prodata/1/aa/6N.COMB.pep.*
17: /cgn2_6/prodata/1/aa/6O.COMB.pep.*
18: /cgn2_6/prodata/1/aa/6P.COMB.pep.*
19: /cgn2_6/prodata/1/aa/6Q.COMB.pep.*
20: /cgn2_6/prodata/1/aa/6R.COMB.pep.*
21: /cgn2_6/prodata/1/aa/6S.COMB.pep.*
22: /cgn2_6/prodata/1/aa/6T.COMB.pep.*
23: /cgn2_6/prodata/1/aa/6U.COMB.pep.*
24: /cgn2_6/prodata/1/aa/6V.COMB.pep.*
25: /cgn2_6/prodata/1/aa/6W.COMB.pep.*
26: /cgn2_6/prodata/1/aa/6X.COMB.pep.*
27: /cgn2_6/prodata/1/aa/6Y.COMB.pep.*
28: /cgn2_6/prodata/1/aa/6Z.COMB.pep.*
29: /cgn2_6/prodata/1/aa/6A.COMB.pep.*
30: /cgn2_6/prodata/1/aa/6B.COMB.pep.*
31: /cgn2_6/prodata/1/aa/6C.COMB.pep.*
32: /cgn2_6/prodata/1/aa/6D.COMB.pep.*
33: /cgn2_6/prodata/1/aa/6E.COMB.pep.*
34: /cgn2_6/prodata/1/aa/6F.COMB.pep.*
35: /cgn2_6/prodata/1/aa/6G.COMB.pep.*
36: /cgn2_6/prodata/1/aa/6H.COMB.pep.*
37: /cgn2_6/prodata/1/aa/6I.COMB.pep.*
38: /cgn2_6/prodata/1/aa/6J.COMB.pep.*
39: /cgn2_6/prodata/1/aa/6K.COMB.pep.*
40: /cgn2_6/prodata/1/aa/6L.COMB.pep.*
41: /cgn2_6/prodata/1/aa/6M.COMB.pep.*
42: /cgn2_6/prodata/1/aa/6N.COMB.pep.*
43: /cgn2_6/prodata/1/aa/6O.COMB.pep.*
44: /cgn2_6/prodata/1/aa/6P.COMB.pep.*
45: /cgn2_6/prodata/1/aa/6Q.COMB.pep.*
46: /cgn2_6/prodata/1/aa/6R.COMB.pep.*
47: /cgn2_6/prodata/1/aa/6S.COMB.pep.*
48: /cgn2_6/prodata/1/aa/6T.COMB.pep.*
49: /cgn2_6/prodata/1/aa/6U.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	86	100.0	339	4	US-09-323-873A-2	Sequence 2, Appl1
2	42.5	49.4	256	4	US-09-134-001C-5278	Sequence 5278, Ap
3	40.5	46.5	221	4	US-09-247-373B-54	Sequence 54, Appl1
4	38	44.2	245	1	US-08-403-866-27	Sequence 2, Appl1
5	38	44.2	245	1	US-08-403-866-27	Sequence 2, Appl1
6	38	44.2	245	1	US-08-403-866-27	Sequence 2, Appl1
7	38	44.2	245	1	US-08-403-866-27	Sequence 2, Appl1
8	38	44.2	245	1	US-08-403-866-27	Sequence 2, Appl1
9	37	43.0	25	1	US-08-468-708B-16	Sequence 16, Appl1
10	37	43.0	25	2	US-08-241-664B-16	Sequence 16, Appl1
11	37	43.0	323	2	US-08-969-106-4	Sequence 4, Appl1
12	37	43.0	416	1	US-08-252-995D-2	Sequence 2, Appl1
13	37	43.0	416	1	US-08-834-108-6	Sequence 2, Appl1
14	37	43.0	416	1	US-08-834-108-6	Sequence 2, Appl1
15	37	43.0	577	1	US-08-484-106-24	Sequence 24, Appl1
16	37	43.0	577	1	US-08-484-106-24	Sequence 24, Appl1
17	37	43.0	577	1	US-08-756-317-13	Sequence 13, Appl1
18	37	43.0	605	4	US-09-394-645-2	Sequence 2, Appl1
19	37	43.0	605	4	US-09-243-560B-2	Sequence 2, Appl1
20	37	43.0	711	4	US-08-961-983-32	Sequence 2, Appl1
21	37	43.0	711	4	US-08-961-983-32	Sequence 2, Appl1
22	37	43.0	752	4	US-09-402-423-6	Sequence 6, Appl1
23	37	43.0	925	1	US-08-255-995D-4	Sequence 4, Appl1
24	37	43.0	925	1	US-08-834-108-4	Sequence 4, Appl1
25	37	43.0	1228	2	US-08-605-785-537	Sequence 537, App
26	37	43.0	1228	2	US-09-439-313-537	Sequence 537, App
27	36	41.9	202	2	US-08-286-819A-6	Sequence 6, Appl1

28	36	41.9	202	3	US-08-980-357-6	Sequence 6, Appl1
29	36	41.9	373	4	US-09-339-401C-1418	Sequence 3, Appl1
30	36	41.9	440	3	US-09-052-778-13	Sequence 13, Appl1
31	36	41.9	449	4	US-09-134-001C-5673	Sequence 5673, Ap
32	36	41.9	550	1	US-08-484-493-2	Sequence 2, Appl1
33	36	41.9	550	1	US-08-484-493-2	Sequence 2, Appl1
34	36	41.9	550	1	US-08-484-493-2	Sequence 2, Appl1
35	36	41.9	550	2	US-08-345-212-2	Sequence 2, Appl1
36	36	41.9	550	4	US-09-249-003-2	Sequence 2, Appl1
37	36	41.9	550	4	US-09-249-003-2	Sequence 2, Appl1
38	36	41.9	550	4	US-09-249-003-2	Sequence 2, Appl1
39	36	41.9	1854	4	US-08-004-838-108	Sequence 108, App
40	36	41.9	2291	2	US-08-286-819A-29	Sequence 29, Appl1
41	36	41.9	2291	3	US-08-980-357-29	Sequence 29, Appl1
42	35.5	41.3	273	4	US-09-651-656-7	Sequence 7, Appl1
43	35.5	41.3	273	4	US-09-650-855-7	Sequence 7, Appl1
44	35.5	41.3	414	5	PCT-US92-06840-2	Sequence 2, Appl1
45	35	40.7	221	1	US-08-484-635-148	Sequence 148, App

ALIGNMENTS

```

RESULT 1
US-09-323-873A-2
Sequence 2, Application US/09323873A
GENERAL INFORMATION:
APPLICANT: Daniel E. Afar
APPLICANT: Rene S. Hubert
APPLICANT: Khan Leong
APPLICANT: Arthur C. Raitano
APPLICANT: Douglas C. Saffran
APPLICANT: Robert S. Thompson
TITLE OR INVENTION: ANTIBODY BINDING TRANSMEMBRANE ANTIGENS
TITLE OR INVENTION: EXPRESSED IN HUMAN CANCERS AND USBS THEREOF
FILE REFERENCE: 129.16US02
CURRENT APPLICATION NUMBER: US/09/323,873A
CURRENT FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: 60/087,520
PRIOR FILING DATE: 1998-06-01
PRIOR APPLICATION NUMBER: 60/091,183
PRIOR FILING DATE: 1998-06-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 339
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-323-873A-2
Query Match 100.0% Score 86; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 9.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 WKPKPRNLEDDYL 15
Db 14 WKPKPRNLEDDYL 28
RESULT 2
US-09-134-001C-5278
Sequence 5278, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Genentech, Inc. et al
TITLE OR INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08

```


Db 293 WKQPSQTEADN 305

```

RESULT 6
US-09-385-752-7
Sequence 7, Application US/09385752
Patent No. 6228591
GENERAL INFORMATION:
APPLICANT: Stefan Somlo and Toshio Mochizuki
TITLE OF INVENTION: POLYCLONAL KIDNEY DISEASE PROZ GENE
TITLE OF INVENTION: POLYCLONAL KIDNEY DISEASE PROZ GENE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSER: AMSTER, ROTHSTEIN & EBENSTEIN
STREET: 90 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE
MEDIUM TYPE: DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
PRIORITY NUMBER: US/09385,752
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/651,999
FILING DATE: MAY 23, 1996
ATTORNEY/AGENT INFORMATION:
NAME: ELIZABETH A. BOGOSIAN
REGISTRATION NUMBER: 39,911
ADDRESS/COMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 968 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: peptide
HYPOTHETICAL: NO
US-09-385-752-7

Query Match 44.2%; Score 38; DB 4; Length 968;
Best Local Similarity 46.2%; Pred. No 2.9e+02;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 WKQPSQTEADN 13
Db 293 WKQPSQTEADN 305

RESULT 7
US-08-061-376-5
Sequence 5, Application US/08061376
Patent No. 6175000
GENERAL INFORMATION:
APPLICANT: Evans, Glen A.
APPLICANT: Shih, Li-Rueh
APPLICANT: Shih, Li-Rueh
APPLICANT: Shih, Li-Rueh
APPLICANT: Shih, Li-Rueh
TITLE OF INVENTION: CHARACTERIZATION OF A CHROMOSOME 11Q23
TITLE OF INVENTION: TRANSLATION BREAKPOINT ASSOCIATED WITH ACUTE LEUKEMIAS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:

```

```

ADDRESSER: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/061,376
FILING DATE: 13-MAY-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
ADDRESS/COMMUNICATION INFORMATION:
TELEPHONE: (619)546-4737
TELEFAX: (619)546-9392
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3969 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-061-376-5

Query Match 44.2%; Score 38; DB 4; Length 3969;
Best Local Similarity 60.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 PRPRLIEDY 14
Db 424 PRPRLIEDY 433

RESULT 8
US-08-468-709B-16
Sequence 5, Application US/08468709B
Patent No. 6554137
GENERAL INFORMATION:
APPLICANT: Astrom, Anders
APPLICANT: Voornhees, John
APPLICANT: Patterson, Ulrika
APPLICANT: Tavaehol, Amir
TITLE OF INVENTION: HUMAN CRAB-I AND CRAB-II
TITLE OF INVENTION: HUMAN CRAB-I AND CRAB-II
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSER: Harnes, Dickey & Pierce, P.L.C.
STREET: PO Box 828
CITY: Bloomfield Hills
STATE: Michigan
COUNTRY: United States of America
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,709B
FILING DATE: 06/06/95
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REGISTRATION NUMBER: 36,683
REFERENCE/DOCKET NUMBER: 2115-00676DVF
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1600

```

TELEFAX: (610) 641-0270
 TELEX: 287637
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 25 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 US-08-468-709B-16

Query Match 43.0%; Score 37; DB 1; Length 25;
 Best Local Similarity 54.5%; Pctd No. 7.9;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 WKRRKRRLE 11
 DB 7 WKRRSENFE 17

RESULT 9
 US-08-241-66AB-16
 Sequence 16, Application US/0824166AB
 Patent No. 5871809

GENERAL INFORMATION:
 APPLICANT: Voothees, John J.
 APPLICANT: Astrom, Andere
 APPLICANT: Paterson, Ulika
 APPLICANT: Paterson, Ulika
 TITLE OF INVENTION: HUMAN CRAB-I AND CRAB-II
 NUMBER OF SEQUENCES: 18
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
 STREET: PO Box 828
 CITY: Bloomfield Hills
 STATE: Michigan
 COUNTRY: United States of America
 ZIP: 48303

COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, version #1.25
 CURRENT APPLICATION DATA: US/08/241,66AB
 FILING DATE: May 11, 1994
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Smith, Deann F.
 REGISTRATION NUMBER: 36, 683
 REFERENCE/DOCKET NUMBER: 2115-00676CDD
 TELEPHONE: (610) 641-1600
 TELEFAX: (610) 641-0270

INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 25 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 US-08-241-66AB-16

Query Match 43.0%; Score 37; DB 2; Length 25;
 Best Local Similarity 54.5%; Pctd No. 7.9;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 WKRRKRRLE 11
 DB 7 WKRRSENFE 17

RESULT 10
 US-08-969-106-4

Sequence 4, Application US/08969106
 Patent No. 5986055
 GENERAL INFORMATION:
 APPLICANT: Nam, M.
 APPLICANT: Namphalan, K.
 APPLICANT: Schulz, V.
 TITLE OF INVENTION: CDK2 INTERACTIONS
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds LLP
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: Patentin Release 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/969,106
 FILING DATE: 13-NOV-1997
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Mirock, S. Leslie
 REGISTRATION NUMBER: 6, 872
 REFERENCE/DOCKET NUMBER: 67934-057
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-790-9090
 TELEFAX: 212-869-9741

INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 323 amino acids
 TYPE: amino acid
 STRANDEDNESS: 3
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-969-106-4

Query Match 43.0%; Score 37; DB 2; Length 323;
 Best Local Similarity 50.0%; Pctd No. 13.4+0.2;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 KRRNLRSDDY 15
 DB 292 KRRGVEDDDY 303

RESULT 11
 US-08-252-995D-2
 Sequence 2, Application US/08252995D
 Patent No. 5650501

GENERAL INFORMATION:
 APPLICANT: Dennis, James W
 APPLICANT: Hefterman, Mike
 APPLICANT: Hefterman, Mike
 TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BERESKIN & PARR
 STREET: 40 King Street West
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5H 3Y2

COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, version #1.30
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/252,995D
FILING DATE: 02-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Kurydyk, Linda M
REGISTRATION NUMBER: 34,971
TELEPHONE/DOCKET NUMBER: 3153-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 364-1398
INVENTOR: DENNIS, JAMES W
SEQUENCE CHARACTERISTICS:
LENGTH: 416 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-252-995D-2

Query Match 43.0%; Score 37; DB 1; Length 416;
Best Local Similarity 54.5%; Pred. No. 1.7e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 4 KPRRLSEDDY 14
|||||:|:|:
Db 406 KPRRLSDENQH 416

RESULT 12
US-08-834-108-2
Sequence 2, Application US/08834108
Patent No. 5976893
GENERAL INFORMATION:
APPLICANT: Dennis, James W
REGISTRATION NUMBER: 34,971
TELEPHONE/DOCKET NUMBER: 3153-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 364-1398
INVENTOR: DENNIS, JAMES W
SEQUENCE CHARACTERISTICS:
LENGTH: 416 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-834-108-2

Query Match 43.0%; Score 37; DB 2; Length 416;
Best Local Similarity 54.5%; Pred. No. 1.7e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 4 KPRRLSEDDY 14

|||||:|:|:
Db 406 KPRRLSDENQH 416

RESULT 13
US-08-252-995D-6
Sequence 6, Application US/08252995D
Patent No. 5650501
GENERAL INFORMATION:
APPLICANT: Dennis, James W
REGISTRATION NUMBER: 34,971
TELEPHONE/DOCKET NUMBER: 3153-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 364-1398
INVENTOR: DENNIS, JAMES W
SEQUENCE CHARACTERISTICS:
LENGTH: 464 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-252-995D-6

Query Match 43.0%; Score 37; DB 1; Length 464;
Best Local Similarity 54.5%; Pred. No. 1.9e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 4 KPRRLSEDDY 14
|||||:|:|:
Db 406 KPRRLSDENQH 416

RESULT 14
US-08-834-108-6
Sequence 6, Application US/08834108
Patent No. 5976893
GENERAL INFORMATION:
APPLICANT: Dennis, James W
REGISTRATION NUMBER: 34,971
TELEPHONE/DOCKET NUMBER: 3153-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 364-1398
INVENTOR: DENNIS, JAMES W
SEQUENCE CHARACTERISTICS:
LENGTH: 416 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-834-108-2

```

/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/   APPLICATION NUMBER: US/08/834,108
/   FILING DATE:
/   CLASSIFICATION: 536
/ ATTORNEY/AGENT INFORMATION:
/   NAME: Rudyguy, Linda M
/   REGISTRATION NUMBER: 34,971
/   REFERENCE/DOCKET NUMBER: 3153-210
/   TELEPHONE: (415) 363-7211
/   TELEFAX: (415) 363-1398
/ INFORMATION FOR SEQ ID NO: 6:
/   SEQUENCE CHARACTERISTICS:
/     LENGTH: 464 amino acids
/     TYPE: amino acid
/     TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-834-108-6

Query Match      43.0%; Score 37; DB 2; Length 464;
Best Local Similarity 54.5%; Pred. No. 1.9e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 KERRNEEDDY 14
DB 406 KPRSLDENQH 416

RESULT 15
US-08-484-105-24
/ Sequence 24, Application US/08484105
/ Patent No. 5589341
/ GENETIC INFORMATION:
/   APPLICANT: STILLMAN, Bruce
/   APPLICANT: BELL, Stephen P
/   APPLICANT: KOBAYASHI, Ryuji
/   APPLICANT: RINE, Jasper
/   APPLICANT: FOSS, Margit
/   APPLICANT: MCNALLY, Francis J
/   APPLICANT: JENSEN, Patricia
/   APPLICANT: HERSHKOWITZ, Mira
/   APPLICANT: LI, Joachim J
/   APPLICANT: GAVIN, Kimberly
/   TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
/   NUMBER OF SEQUENCES: 24
/   CORRESPONDENCE ADDRESS:
/     ADDRESSEE: FLSHR, HOMBACH, TEST, ABRITTON & HENBERT
/     ADDRESS: 3000 Serrano Center, Suite 3400
/     CITY: San Francisco
/     STATE: California
/     COUNTRY: USA
/     ZIP: 94111-4187
/ COMPUTER READABLE FORM:
/   MEDIUM TYPE: floppy disk
/   COMPUTER: IBM PC compatible
/   OPERATING SYSTEM: PC-DOS/MS-DOS
/   SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/   APPLICATION NUMBER: US/08/484,105
/   FILING DATE:
/   CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/   NAME: Omen P.L.L.C., Richard Aton
/   REGISTRATION NUMBER: 34,971
/   REFERENCE/DOCKET NUMBER: A-59032/DJB/BAO
/ TELECOMMUNICATION INFORMATION:
/   TELEPHONE: (415) 494-8700
/   TELEFAX: (415) 494-8771
/   TELEX: 910 277299
/ INFORMATION FOR SEQ ID NO: 24:

```

```

/ SEQUENCE CHARACTERISTICS:
/   LENGTH: 577 amino acids
/   TYPE: amino acid
/   TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-484-105-24

Query Match      43.0%; Score 37; DB 1; Length 577;
Best Local Similarity 70.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 KPRRNLEEDD 13
DB 52 KPRYLEEDD 61

Search completed: March 26, 2003, 16:51:37
Job time : 11.593 secs

```


PA (LEON/) LEONG K.
PA (SALT/) RATTANO A. B.
PA (SAFE/) SAEFRAN D. C.
XX
XX
PI Afar DE, Hubert RS, Leong K, Raitano AB, Saefran DC
XX
WR; 2000-072832/06.

PT Novel proteins useful as diagnostic markers and therapeutic targets,
 PT particularly for prostatic cancer -
 XX
 PS Example 3C, Page 41, 83pp, English.

This sequence represents a synthetic peptide (peptide 1) corresponding to residues 14-28 of STRAP-1 (seipentine transmembrane antigen of the prostate, AAV58194). This peptide was used to raise the polyclonal anti-STRAP-1 N-terminus antibodies in sheep. The antibodies were used to probe Western blot of cell lysates from prostate and non-prostate cancer cell lines, and in immunohistochemical studies of paraffin-embedded tissue sections. STRAP-1 is a member of the STRAP family of proteins (AAV58194-758197) which exhibit a high degree of structural conservation, but which show no significant structural homology to known human proteins. The STRAP-1 gene has been localised to chromosome 7p22. STRAP-1 is thought to be a type IIIa membrane protein and is expressed predominantly in prostate cells in normal human tissues. Structurally, STRAP-1 is a 339 amino acid protein characterised by six transmembrane domains and an intracellular N- and C-termini, suggesting that it could function as a multi-pass transmembrane protein. The expression of STRAP-1 mRNA and protein expression is maintained at high levels and throughout all stages of prostate cancer. STRAP-1 mRNA and/or protein is also overexpressed in certain other cancers, including bladder, colon, pancreatic and ovarian cancer. The function of the STRAP proteins is not known. They may be ion channels (from the presence of six transmembrane domains), a feature which is shared by certain ion channels) or gap-junction proteins (from immunohistochemical staining). STRAP-1 and STRAP-2 are cell-surface tumour antigen immunisation targets. STRAP-1 STRAP-expressing cells STRAP proteins may be used to identify specific-binding agents, to produce anticancer vaccines and to generate specific antibodies. The antibodies may be used for detection, prognostic and monitoring of cancers (or susceptibility to cancer), as therapeutic inhibitors or to target therapeutic agents to their site of action. STRAP inhibitors or to target therapeutic agents to their site of action. STRAP antibodies and prognostic reagents, for identifying STRAP-expressing cells for screening in drug screens. High levels of STRAP proteins are exposed on the cell surface. They are easily targeted by systemically administered agents, and because they are expressed mainly on prostatic epithelial cells, agents targeted to them should have minimal side effects on other tissues.

SQ Sequence 15 AA;

Query Match	Score	DB	Length
100.0%	86	21	15

Best Local Similarity 100.0%; Pred. No. 4.6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	WKMKPRNLEBDDYL	15
Db	1	WKMKPRNLEBDDYL	15

RESULT 2	
AAE02784	
ID	AAE02784 standard; peptide; 15 AA

AC AAE02784 ;

DT 06-AUG-2001 (first entry)

DE Immunogenic peptide #1 of human STEAP-1.

KM	Human; cytosolic; antiproliferative; vaccine; gene therapy;
KM	Hixn transmembrane epithelial antigen of the prostate-1; STEAP-1;
KM	human chromosome 7p22.3; cancer; prostate; colon; bladder; pancreatic;
KM	ovarian; lung; serpentine transmembrane antigen; immunogen.
XX	
OS	Homo sapiens.

PN WO200140276-A2.

PD 07-JUN-2001.

PF 06-DEC-2000; 2000MO-US33040.

PR 06-DEC-1999;

PA (UROG-) UROGENI

PI Afar DEH, H

XX

XX

PT proteins, expressed in

✕✕

XX

CC The present invention relates to human *ds* transmembrane epithelial
CC antigen of the prostate (STEP) protein. STEP gene is a member of cell
CC surface receptor transmembrane antigens. STEP gene is used in gene
CC therapy, inhibiting the development or progression of a cancerous
CC cell, or inhibiting growth or killing cells expressing STEP in a patient,
CC comprising administering a vaccine composition to the patient. Treating
CC a patient with a cancer that expresses STEP, or inhibiting growth or
CC killing cells expressing STEP, comprises administering to the patient a
CC vector encoding single chain monoclonal antibody that comprises the
CC variable domains of the heavy and light chains of the monoclonal antibody
CC that specifically binds to STEP, such that the vector delivers the
CC single chain monoclonal antibody coding sequence expressed
CC intracellularly. The present sequence is an immunogenic peptide of
CC STEP-1. STEP-1 gene is located on chromosome 7p23.3. This peptide is
CC used to immunize sheep for the generation of sheep polyclonal antibodies
CC towards the amino-terminus of anti-STEP-1.

SQ Sequence 15 AA;

Query Match 100.0%; Score 86; DB 22; Length 15

Best Local Similarity 100.0%; Pred. No. 4,6e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```
QY      1 WKMKPRRLLEEDYL 15
          |||||
Db      1 WKMKPRRLLEEDYL 15
```

RESULT	3
AA12305	
ID	AA12305 standard; Protein; 66 AA

AC MAY12305;

DT 17-JUN-1999 (first entry)

Human 5' EST secreted protein SEQ ID NO:336.

KM Human, secreted protein; EST, expressed sequence tag; diagnosis;
 KM forensic; gene therapy; chromosome mapping; signal peptide;
 KM upstream regulatory sequence; cytokine activity; cell proliferation;
 KM differentiation; haemopoietic regulation; tissue growth regulation;
 KM reproductive hormone regulation; chemotactic; chemokinetic; hematopoietic

KW Thrombolytic; anti-inflammatory; tumour inhibition.
 XX Homo sapiens.
 OS MO3906548-42.
 PM 11-FEB-1999.
 PD 11-FEB-1999.
 XX 31-JUL-1998; 98MO-1B01222.
 PF 01-AUG-1997; 97US-0905135.
 PR (GIST) GENSET.
 XX Ductect A, Dumas Milne Edwards J, Lacroix B;
 PI MPI; 1999-153778/13.
 XX N-PSDB; AAX41138.
 DR New nucleic acids encoding human secreted proteins - obtained from
 PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
 PT kidney, lung, umbilical cord, placenta and colon tissue
 XX Claim 27; Page 677-678; 824pp; English.
 PS
 XX AAX41099 to AAX41347 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAX1261 to
 CC AAX1284, respectively. The proteins given represent the signal peptide
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, hematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
 CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.
 XX Sequence 66 AA;
 SQ
 Query Match 100.0%; Score 86; DB 20; Length 66;
 Best Local Similarity 100.0%; Pred. No. 2,3e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WKMKPRNLSEDDYL 15
 DB 14 WKMKPRNLSEDDYL 28
 RESULT 4
 ID AAY12304 standard; Protein; 109 AA.
 AC AAY12304;
 XX 17-JUN-1999 (first entry)
 DT Human 5' EST secreted protein SBQ ID NO.335.
 XX Human 5' EST secreted protein SBQ ID NO.335.
 DE Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequences; cytochrome P-450; secretion;
 KW effect; tumour inhibition; tissue growth; cell proliferation;
 KW reproductive hormone regulation; chemotactic/chemokinetic; haemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition.
 XX Homo sapiens.
 OS

PM MO3906548-42.
 XX 11-FEB-1999.
 OS 31-JUL-1998; 98MO-1B01222.
 XX 01-AUG-1997; 97US-0905135.
 PR (GIST) GENSET.
 XX Ductect A, Dumas Milne Edwards J, Lacroix B;
 PI MPI; 1999-153778/13.
 XX N-PSDB; AAX41137.
 DR New nucleic acids encoding human secreted proteins - obtained from
 PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
 PT kidney, lung, umbilical cord, placenta and colon tissue
 XX Claim 27; Page 677; 824pp; English.
 PS
 XX AAX41099 to AAX41347 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAX1261 to
 CC AAX1284, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, hematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, receptor/ligand activity, anti-inflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.
 XX Sequence 109 AA;
 SQ
 Query Match 100.0%; Score 86; DB 20; Length 109;
 Best Local Similarity 100.0%; Pred. No. 3,6e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WKMKPRNLSEDDYL 15
 DB 14 WKMKPRNLSEDDYL 28
 RESULT 5
 ID AAY58194 standard; Protein; 339 AA.
 AC AAY58194;
 XX 14-MAR-2000 (first entry)
 DT Human STRAP-1 protein.
 XX Serpentine transmembrane antigen of the prostate; STRAP-1; prostate;
 KW transmembrane domain; type IIIa membrane protein; expression; cancer;
 KW prostate cancer; bladder cancer; colon cancer; pancreatic cancer;
 KW ovarian cancer; tumour antigen; immunisation; immune response;
 KW cellular; humoral; anticancer vaccine; antibody; detection; diagnosis;
 KW prognosis; monitoring; susceptibility; therapeutic inhibitor;
 KW drug targeting; recombinant protein.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH 1..69
 FT Region /note= "intracellular region 1"

FT	Domain	70..91
FT	Region	/note="Transmembrane domain 1"
FT	Region	92..113 /note="extracellular region 1 (AA158198)"
FT	Domain	114..135 /note="Transmembrane domain 2"
FT	Region	137..162 /note="intracellular region 2"
FT	Domain	163..184 /note="Transmembrane domain 3"
FT	Region	185..218 /note="extracellular region 3 (AA158199)"
FT	Domain	219..243 /note="Transmembrane domain 4"
FT	Region	243..253 /note="intracellular region 3"
FT	Domain	252..276 /note="Transmembrane domain 5"
FT	Region	277..291 /note="extracellular region 3 (AA158200)"
FT	Domain	292..313 /note="Transmembrane domain 6"
FT	Region	314..339 /note="intracellular region 4"
FT	Region	
XX	MO962941-A2.	
XX	09-DEC-1999.	
XX	01-JUN-1999;	99NC-US12157.
XX	01-JUN-1998;	98US-0087520.
XX	30-JUN-1998;	98US-0091183.
XX	(UNOC-) UNOGENESIS INC.	
XX	(HUBER-) HUBER R. S.	
XX	(LEONG-) LEONG K.	
XX	(RAIT-) RAITANO A. B.	
XX	(SAFF-) SAFFRAN D. C.	
XX	Afar DE, Hubert RS, Leong K, Raitano AB, Safran DC;	
XX	WPI. 2000-072832/05.	
XX	N-PSDB; AAZ49395; AAZ49396.	
XX	Novel proteins useful as diagnostic markers and therapeutic targets,	
XX	particularly for prostatic cancer -	
XX	Claim 1: Fig 1A: 83pp: English.	

CC specific antibodies. The antibodies may be used for detection, prognosis
CC and monitoring of cancers (or susceptibility to cancer), as therapeutic
CC inhibitors or to target therapeutic agents to their site of action. STRAP
CC nuclear acids may be used for recombinant protein production, as
CC probes for screening inhibitors of STRAP expression and for therapeutic
CC modulation/inhibition of STRAP expression, since high levels of STRAP
CC proteins are exposed on the cell surface, they are easily targeted by
CC systemically administered agents, and because they are expressed mainly
CC on prostatic epithelial cells, agents targeted to them should have
CC minimal side effects on other tissues.

DB Db

QY Sequence 339 AA:
1 WKAKPRNLLEDDYL 15
14 WKAKPRNLLEDDYL 28

RESULT 6
AAUS6927
ID AAUS6927 standard; Protein; 339 AA.
XX AAUS6927;
DT 30-JAN-2002 (first entry)
DE Human prostate cDNA encoded protein #72.
XX Human prostate cancer; cytostatic; immunostimulant; tumour; immunogen.
XX Homo sapiens.
XX Homo sapiens.
XX MO200173032-AZ.
PD 04-OCT-2001.
XX 27-MAR-2001; 2001NC-USO9919.
PF 27-MAR-2001; 2000US-0536657.
PI 05-MAY-2001; 2000US-0558100.
PR 12-MAY-2001; 2000US-0570737.
PR 13-JUN-2001; 2000US-0593793.
PR 27-JUN-2001; 2000US-0605783.
PR 10-AUG-2001; 2000US-0636215.
PR 20-AUG-2001; 2000US-0651236.
PR 06-SEP-2001; 2000US-0674219.
PR 10-OCT-2001; 2000US-0692429.
PR 10-OCT-2001; 2000US-0695166.
XX (CORI-) CORIXA CORP.
PA Xu J, Dillon DC, Mitcham JT, Harlocker SU, Jiang Y, Kalos MD,
PI Penger GR, Retter WJ, Stoick JA, Day CH, Vadvick TS, Carter D,
PI Li SX, Wang A, Skeiky TAW, Hepler WT, Henderson RH,
DR NPSB; 2001-639323/77.
XX N-PSDB; AAS64160.
PT New human prostate-specific polypeptides and polymucleotides useful for
PT the diagnosis and treatment of cancer, especially prostate cancer -
XX Claim 2; Page 549; 57pp; English.
XX The invention relates to isolated prostate-specific
CC polymucleotides, polypeptides, fusion proteins of the polypeptides,
CC antibodies raised against the polypeptides (or antigenic epitopes
CC derived from them) and antigen-presenting cells expressing the
CC polypeptides. The antibodies are useful for detecting the presence of

CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
 CC the antigen-presenting cells are useful for stimulating and/or expanding
 CC T cells, especially cytotoxic T cells, for cancer therapy. The polypeptides
 CC of cancer specifically protect, compositions comprising the
 CC polynucleotide and/or polypeptide are useful for stimulating an immune
 CC response, and for treating cancer. The oligonucleotide is useful for
 CC detecting cancer. The present sequence is a prostate specific
 CC polypeptide of the invention.

XX Sequence 339 AA;

Seq Match 100.0%; Score 86; DB 22; Length 339;

Best Local Similarity 100.0%; Pred. No. 1.3e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WQMKPRNLEEDDTL 15

DB 14 WQMKPRNLEEDDTL 28

RESULT 7

AAW78845 standard; Protein; 339 AA.

AAW78845;

06-NOV-2001 (first entry)

Human protein SEQ ID NO 1507.

Human; cytokine; cell proliferation; cell differentiation; gene therapy;

vacine; peptide therapy; stem cell growth factor; haematopoiesis;

tissue growth factor; immunomodulatory; cancer; leukemia;

nervous system disorder; arthritis; inflammation.

Homo sapiens.

MO200157190-A2.

09-AUG-2001.

05-FEB-2001; 2001MO-US04098.

01-FEB-2001; 2000US-0486314.

27-APR-2001; 2000US-0560875.

20-JUN-2001; 2000US-0598075.

19-JUL-2001; 2000US-0620325.

01-SEP-2001; 2000US-0654936.

15-SEP-2001; 2000US-0663561.

20-OCT-2001; 2000US-0693325.

30-NOV-2001; 2000US-0728422.

(HYSB-) HYSB INC.

Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y,

Zhao Q, Wang D, Zhang J, Ren F, Chen R, Wang ZW,

Xue A, Yang Y, Wejherman T, Goodrich R;

NPI; 2001-476283/51.

N-PDB; JAKS1978.

Nucleic acids encoding polypeptides with cytokine-like activities,

useful in diagnosis and gene therapy -

CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibin activity and may be useful in the diagnosis and
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and
 CC inflammation. cancer, leukemia, nervous system disorders, arthritis and
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAH80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.

XX Sequence 339 AA;

Seq Match 100.0%; Score 86; DB 22; Length 339;

Best Local Similarity 100.0%; Pred. No. 1.3e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WQMKPRNLEEDDTL 15

DB 14 WQMKPRNLEEDDTL 28

RESULT 8

AAW01282 standard; Protein; 339 AA.

AAW01282;

04-OCT-2001 (first entry)

P789P amino acid sequence.

Human; prostate cancer; prostate-specific; diagnosis; vaccine;

cytostatic; gene therapy; metastasis.

Homo sapiens.

MO200151633-A2.

19-JUL-2001.

16-JAN-2001; 2001MO-US01574.

14-JAN-2001; 2000US-0483672.

(CORI-) CORIXA CORP.

Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG,

Kales ML, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW,

WPI; 2001-425873/45.

New polynucleotide encoding a prostate-specific protein, for

diagnosing, monitoring and treating prostate cancer in a patient and

for use in vaccine -

Claim 2; Page 510-512; 543PP; English.

The present invention describes polynucleotide sequences (I) which encode

prostate-specific protein, the production and gene therapy (II) (III)

antibodies to (II), fusion proteins comprising (II) and isolated

T cells prepared using (I) or (II) are used to treat cancer in a patient.

(I) and the antibodies are also used in the detection of cancer in a

patient. The cancer that is diagnosed or treated is particularly

prostate cancer. (I) and (II) can be used in vaccines. The antibodies or

(I) can be used for monitoring the progression of cancer in a patient.

Methods for diagnosing, monitoring and treating prostate cancer, as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to

AAH01313 represent polynucleotide and amino acid sequences used in the

embodiment of the present invention.

Sequence 339 AA;

PI LI SX, Skeiky YAM, Hepler WT, Henderson RA;
 XX WPI: 2002-255649/30.
 DR New prostate-specific polynucleotides for diagnosing and treating
 PT diseases, in particular prostate cancer, and as markers for the
 PT progression of cancer -
 XX
 PS Claim 2; SEQ ID NO 879; 879p; English.
 CC The present invention provides prostate-specific coding sequences and
 CC nucleic acids encoding polypeptides of the present invention
 CC of cancers, particularly prostate cancer. The present sequence is a
 CC protein described in the invention.
 XX
 SO Sequence 339 AA;
 Query Match 100.0%; Score 86; DB 23; Length 339;
 Best Local Similarity 100.0%; Pred. No. 1,3e-08;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 WKKKPRNLEEDYL 15
 DB 14 WKKKPRNLEEDYL 28
 RESULT 11
 ID AAM79829
 AC AAM79829 standard; Protein; 374 AA.
 XX
 XX AAM79829;
 DT 06-NOV-2001 (first entry)
 XX
 XX Human protein SEQ ID NO 3475.
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 OS Homo sapiens.
 XX
 XX NO200157190-A2.
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001MO-US04098.
 XX
 PR 03-FEB-2000; 2000US-0486914.
 PR 27-MAR-2000; 2000US-0588125.
 PR 20-JUN-2000; 2000US-0588125.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0653561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 XX (HSGE-) HSGE INC.
 XX
 PI Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QY, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI: 2001-476283/51.
 DR N-PSDB; AAK52962.
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX
 PS Claim 20; Page 350; 6221pp; English.
 XX
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the

CC encoded polypeptides (AAM79333-AAK80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of cytokines and cell proliferation or which may induce
 CC production of cytokines and cell proliferation are useful in gene therapy or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC The invention provides for SEQ ID NO 3110 (AAK52881), 2111 (AAK52882) and 3666
 CC (AAK80302) as the relevant sequences from the sequence listing
 CC were missing at the time of publication.
 XX
 SO Sequence 374 AA;
 Query Match 100.0%; Score 86; DB 22; Length 374;
 Best Local Similarity 100.0%; Pred. No. 1,4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 WKKKPRNLEEDYL 15
 DB 49 WKKKPRNLEEDYL 63
 RESULT 12
 ID AAB02780
 AC AAB02780 standard; Protein; 375 AA.
 XX
 XX AAB02780;
 DT 06-AUG-2001 (first entry)
 XX
 XX Human six transmembrane epithelial antigen of prostate (STAMP)-1 protein.
 KW Human; cystic; antiproliferative; vaccine; gene therapy;
 KW six transmembrane epithelial antigen of the prostate-1; STAMP-1;
 KW chromosome 7p22.3; cancer; prostate; colon; bladder; lung; ovarian;
 KW pancreatic.
 OS Homo sapiens.
 XX
 XX Key
 XX Location/Qualifiers
 XX 14..28
 XX /label= Immunogenic_peptide #1
 XX 70..91
 XX /label= Transmembrane_domain #1
 XX 86..94
 XX /label= HLA-A2 binding_peptide #2
 XX 111..126
 XX /label= Transmembrane_domain #2
 XX 158..166
 XX /label= HLA-A2 binding_peptide #5
 XX 163..184
 XX /label= Transmembrane_domain #3
 XX 165..173
 XX /label= HLA-A2 binding_peptide #1
 XX 219..241
 XX /label= Transmembrane_domain #4
 XX 254..276
 XX /label= Transmembrane_domain #5
 XX 262..270
 XX /label= HLA-A2 binding_peptide #3
 XX 292..313
 XX /label= Transmembrane_domain #6
 XX 315..310
 XX /label= HLA-A2 binding_peptide #4
 XX 319..340
 XX Misc-difference 339..340
 XX /note= "Encoded by TTGTGCAAT"
 XX
 XX MO200140276-A2.
 XX
 XX 07-JUN-2001.

XX 06-DEC-2000; 2000MO-US33040.
 XX 06-DEC-1999; 99US-0455486.
 XX (UROG-) GENESYS INC.
 XX Afar DEH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC, Faris M;
 XX Jakobovits A;
 XX WPI, 2001-367804/38.
 XX N-PDB; AAD07067.
 XX New STEP (six transmembrane epithelial antigen of the prostate)
 XX PT protein, expressed in human cancers, useful for detecting and treating
 XX cancer -
 XX Example 2, Fig 1A-1B, 18Pp; English.
 CC The present sequence is human six transmembrane epithelial antigen of
 CC the prostate (STEP)-1 protein of clone 10. STEP-1 gene is located on
 CC surface serpinin transmembrane antigens. STEP-1 gene is located on
 CC chromosome 7p22.3 and is used in gene therapy. Inhibiting the development
 CC or progression of a cancer (eg. prostate, colon, bladder, lung, ovarian
 CC and pancreatic) expressing STEP or inhibiting growth or killing cells
 CC expressing STEP in a patient suppresses tumour growth and cancer that
 CC expressing STEP in a patient suppresses tumour growth and cancer that
 CC suppresses STEP, or inhibiting growth or killing cells expressing STEP,
 CC commercial antibodies that comprise the variable domains of the heavy and
 CC monoclonal antibody that comprises the variable domains of the heavy and
 CC light chain of the monoclonal antibody that specifically binds to STEP,
 CC such that the vector delivers the single chain monoclonal antibody coding
 CC sequence to the cancer cells and the encoded single chain monoclonal
 CC antibody binds to STEP and inhibits growth of the cancer cells.
 CC Note: The present sequence is also shown in sequence listing of the
 CC specification, but it lacks amino acid residues at its N-terminal end.
 XX
 XX Sequence 375 AA;
 XX
 XX Query Match 100.0%; Score 86; DB 22; Length 375;
 XX Best Local Similarity 100.0%; Pred. No. 1, 4e-06;
 XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX DB 14 WKPKPRRLLEDDYL 15
 XX 1 WKPKPRRLLEDDYL 15
 XX 14 WKPKPRRLLEDDYL 28
 XX
 XX RESULT 13
 XX AA011840
 XX ID AA011840 standard; Protein; 95 AA.
 XX
 XX AA011840;
 XX
 XX 18-JUN-1999 (first entry)
 XX DT Human 5' EST secreted protein SEQ ID No: 440.
 XX
 XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
 XX Forensic; gene therapy; chromosome mapping; signal peptide; prostate;
 XX upstream regulatory sequence; cytokine activity; cell proliferation;
 XX differentiation; haematopoiesis regulation; tissue growth regulation;
 XX reproductive hormone regulation; chemocarcin; chemocarcin; haematologic;
 XX thrombolytic; anti-inflammatory; tumour inhibition.
 XX OS Homo sapiens.
 XX PN M09306550-A2.
 XX PD 11-FEB-1999.
 XX XX 31-JUL-1998; 98MO-1B01232.
 XX

XX 01-AUG-1997; 97US-0905144.
 XX (BEST) GENSET.
 XX Duclet A, Dumas Milne Edwards J, Lacroix B;
 XX WPI, 1999-13780/13.
 XX DR N-PDB; AAX40562.
 XX New isolated prostate-derived nucleic acids - used to develop
 XX PT prostate-specific antigen immunoregulatory haematopoiesis
 XX PT regulating, anti-inflammatory or tumour inhibition activity
 XX
 XX Claim 34; Page 577; 675p; English.
 XX
 XX AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins expressed in prostate, and encode the protein
 CC given in AAX40438 to AAX40715. The present sequence is a fragment of
 CC the nucleic acid sequence and an N-terminal fragment of secreted protein. The
 CC nucleic acid sequences can be used for developing products for diagnosis and
 CC therapy. The proteins obtained may have cytokine activity, cell
 CC proliferation and differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, chemocarcin/chemocarcin activity, haematologic and
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptides can be used for
 CC polypeptide extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.
 XX
 XX Sequence 95 AA;
 XX
 XX Query Match 81.4%; Score 70; DB 20; Length 95;
 XX Best Local Similarity 100.0%; Pred. No. 0.0002;
 XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX DB 3 MKRRNLEDDYL 15
 XX 1 MKRRNLEDDYL 13
 XX
 XX RESULT 14
 XX AA06309
 XX ID AA06309 standard; Protein; 339 AA.
 XX
 XX AA06309;
 XX
 XX 01-MAR-1999 (first entry)
 XX DT Kidney injury associated molecule HM018 protein.
 XX
 XX Kidney injury associated molecule; kidney injury related molecule;
 XX KIM; tissue growth promotion; regeneration; renal condition;
 XX acute renal failure; acute nephritis; tumour.
 XX OS Rattus sp.
 XX PN M0963071-A1.
 XX PD 26-NOV-1998.
 XX XX 22-MAY-1998; 98MO-US10547.
 XX PR 23-MAY-1997; 97US-0047491.
 XX PR 23-MAY-1997; 97US-0047490.
 XX (BIOJ) BIOGEN INC.
 XX PA Cate RL, Hession CA, Sanicola-Nadel M, Wei H;
 XX PT


```

Query Match          100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 9,5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WKKKPRRLLEDDYL 15
Db 14 WKKKPRRLLEDDYL 28

```

```

RESULT 2
US-09-895-793-879
/ Sequence 879; Application US/09895793
/ Publication No. US20020187472A1
/ GENERAL INFORMATION:
/ APPLICANT: Lai, Preeti
/ APPLICANT: Paris, Mary
/ APPLICANT: Chen, Hui-Mei
/ APPLICANT: Zhen, Craig H.
/ TITLE OF INVENTION: HSP70-RELATED PROTEIN
/ FILE REFERENCE: RC-0037 US
/ CURRENT FILING DATE: 2001-03-09
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: PERL Program
/ SEQ ID NO 11
/ LENGTH: 339
/ ORGANISM: Homo sapiens
/ FEATURE: misc feature
/ NAME/KEY: name feature
/ OTHER INFORMATION: Inocyte ID No. US20020187472A1 96572948
US-09-802-520-11

```

```

Query Match          100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 9,5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WKKKPRRLLEDDYL 15
Db 14 WKKKPRRLLEDDYL 28

```

```

RESULT 3
US-09-895-793-879
/ Sequence 879; Application US/09895793
/ Publication No. US20020192763A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Dillon, David C.
/ APPLICANT: Harlocker, Susan L.
/ APPLICANT: Jilang, Yuguu
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darick
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skelley, Yashir A.W.
/ APPLICANT: Hepler, William T.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Hural, John
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Foy, Teresa
/ APPLICANT: Vinals de Basols, Carlotia
/ APPLICANT: Fanger, Gary R.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.534CZ
/ CURRENT APPLICATION NUMBER: US/09/895.793

```

```

/ CURRENT FILING DATE: 2001-06-29
/ NUMBER OF SEQ ID NOS: 982
/ SOFTWARE: FASTSeq for Windows Version 3.0
/ SEQ ID NO 879
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-895-793-879

```

```

Query Match          100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 9,5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WKKKPRRLLEDDYL 15
Db 14 WKKKPRRLLEDDYL 28

```

```

RESULT 4
US-09-895-814-879
/ Sequence 879; Application US/09895814
/ Publication No. US2002019326A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Dillon, David C.
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: Harlocker, Susan L.
/ APPLICANT: Jilang, Yuguu
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darick
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skelley, Yashir A.W.
/ APPLICANT: Hepler, William T.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Hural, John
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Houghton, Raymond L.
/ APPLICANT: Vinals de Basols, Carlotia
/ APPLICANT: Foy, Teresa
/ APPLICANT: Fanger, Gary R.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.437C26
/ CURRENT APPLICATION NUMBER: US/09/895.814
/ NUMBER OF SEQ ID NOS: 990
/ SOFTWARE: FASTSeq for Windows Version 3.0
/ SEQ ID NO 879
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-895-814-879

```

```

Query Match          100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 9,5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WKKKPRRLLEDDYL 15
Db 14 WKKKPRRLLEDDYL 28

```

```

RESULT 5
US-10-011-095-2
/ Sequence 2; Application US/10011095
/ Publication No. US20030045682A1
/ GENERAL INFORMATION:
/ APPLICANT: Alar, Daniel

```

```

/ APPLICANT: Hubert, Rene S.
/ APPLICANT: Leong, Kahn
/ APPLICANT: Leong, Kahn B.
/ APPLICANT: Safir, Michael C.
/ APPLICANT: Mitchell, Steve Chappell
/ TITLE OF INVENTION: ANTIBODIES IMMUNOSPECIFIC FOR STEAP1 (AS AMENDED)
/ FILE REFERENCE: 511582001610
/ CURRENT APPLICATION NUMBER: US/10/011.095
/ CURRENT FILING DATE: 2001-12-06
/ PRIOR APPLICATION NUMBER: 09/331.873
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/087.520
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/091.183
/ PRIOR FILING DATE: 1998-06-30
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ SEQ ID NO 339
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: DNA
US-10-011-095-2

```

```

Query Match      100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 9.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 WKMKPRRNLEEDDTL 15
DB      14 WKMKPRRNLEEDDTL 28

```

```

RESULT 6
US-10-010-667a-2
/ Sequence 2. Application US/10010667A
/ Publication No. US20030055217A1
/ GENERAL INFORMATION:
/ APPLICANT: Afar, Daniel
/ APPLICANT: Hubert, Rene S.
/ APPLICANT: Leong, Kahn
/ APPLICANT: Leong, Kahn B.
/ APPLICANT: Safir, Michael C.
/ APPLICANT: Mitchell, Steve Chappell
/ TITLE OF INVENTION: NOVEL SERENITIN TRANSMEMBRANE ANTIGENS
/ FILE REFERENCE: 511582001601
/ CURRENT APPLICATION NUMBER: US/10/010.667A
/ CURRENT FILING DATE: 2001-12-06
/ PRIOR APPLICATION NUMBER: 09/331.873
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/087.520
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/091.183
/ PRIOR FILING DATE: 1998-06-30
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ SEQ ID NO 339
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-010-667a-2

```

```

Query Match      100.0%; Score 86; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 9.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 WKMKPRRNLEEDDTL 15
DB      14 WKMKPRRNLEEDDTL 28

```

```

RESULT 7
US-09-759-143-879
/ Sequence 879. Application US/09759143
/ Patent No. US20020022248A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jianshun
/ APPLICANT: Dillon, David C.
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: Harlocker, Susan L.
/ APPLICANT: Jiang, Yuqi
/ APPLICANT: Jiang, Yuqi
/ APPLICANT: Carter, Darick A.
/ APPLICANT: Kales, Michael D.
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darick
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skeiky, Yaser A.W.
/ APPLICANT: Hegler, William
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.427C23
/ CURRENT APPLICATION NUMBER: US/09/759.143
/ CURRENT FILING DATE: 2000-01-12
/ NUMBER OF SEQ ID NOS: 934
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 879
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-759-143-879

```

```

Query Match      100.0%; Score 86; DB 10; Length 339;
Best Local Similarity 100.0%; Pred. No. 9.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 WKMKPRRNLEEDDTL 15
DB      14 WKMKPRRNLEEDDTL 28

```

```

RESULT 8
US-09-780-669-879
/ Sequence 879. Application US/09780669
/ Patent No. US20020051977A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jianshun
/ APPLICANT: Harlocker, Susan L.
/ APPLICANT: Jiang, Yuqi
/ APPLICANT: Harlocker, Susan L.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Kales, Michael D.
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Skeiky, Yaser A.W.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darick
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skeiky, Yaser A.W.
/ APPLICANT: Hegler, William
/ APPLICANT: Mitchell, Patricia D.
/ APPLICANT: Houghton, Raymond L.
/ TITLE OF INVENTION: COMPOSITIONS OF PROSTATE CANCER
/ FILE REFERENCE: 210121.427C24
/ CURRENT APPLICATION NUMBER: US/09/780.669

```

```

/ CURRENT FILING DATE: 2001-02-09
/ NUMBER OF SEQ ID NOS: 943
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 1
/ SEQ ID NO 319
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-09-780-663-879

Query Match      100.0%; Score 86; DB 10; Length 339;
Best Local Similarity 100.0%; Pred. No. 9.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WKKKPRNLEEDYL 15
      |||||
Db      14 WKKKPRNLEEDYL 28

RESULT 9
US-09-823-827-879
/ Sequence 879; Application US/09822827
/ Patent No. US2002008180A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ PREVENTION OF PROSTATE CANCER
/ CURRENT FILING DATE: 2001-03-28; US/09/822,827
/ NUMBER OF SEQ ID NOS: 982
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 879
/ LENGTH: 339
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-09-822-827-879

Query Match      100.0%; Score 86; DB 10; Length 339;
Best Local Similarity 100.0%; Pred. No. 9.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WKKKPRNLEEDYL 15
      |||||
Db      14 WKKKPRNLEEDYL 28

RESULT 10
US-09-801-574-32
/ Sequence 32; Application US/09801574
/ Patent No. US20020081592A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Peijong Jeremy
/ TITLE OF INVENTION: Reproduction-Specific Genes
/ FILE REFERENCE: 0399,2007-002
/ CURRENT FILING DATE: 2001-03-07; US/09/801,574
/ NUMBER OF SEQ ID NOS: 943
/ PRIOR FILING DATE: 2000-01-07; 0/187,518
/ PRIOR APPLICATION NUMBER: 60/261,557
/ NUMBER OF SEQ ID NOS: 90
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 12
/ LENGTH: 1018
/ TYPE: PRF
/ ORGANISM: Mus musculus
US-09-801-574-32

Query Match      48.8%; Score 42; DB 10; Length 1018;
Best Local Similarity 53.8%; Pred. No. 73;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

```

QY      1 WKKKPRNLEED 13
      |||
Db      646 WKLVKSNLESD 658

RESULT 11
US-09-801-574-76
/ Sequence 76; Application US/09801574
/ Patent No. US20020081592A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Peijong Jeremy
/ TITLE OF INVENTION: Reproduction-Specific Genes
/ FILE REFERENCE: 0399,2007-002
/ CURRENT FILING DATE: 2001-03-07; US/09/801,574
/ NUMBER OF SEQ ID NOS: 943
/ PRIOR FILING DATE: 2000-01-07; 0/187,518
/ PRIOR APPLICATION NUMBER: 60/261,557
/ NUMBER OF SEQ ID NOS: 90
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 76
/ LENGTH: 1019
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-09-801-574-76

Query Match      48.8%; Score 42; DB 10; Length 1019;
Best Local Similarity 53.8%; Pred. No. 74;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      1 WKKKPRNLEED 13
      |||
Db      645 WKLVKSNLESD 657

RESULT 12
US-09-813-453A-45
/ Sequence 45; Application US/09813453A
/ Patent No. US20020168681A1
/ GENERAL INFORMATION:
/ APPLICANT: Paterson, Thomas A.
/ TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
/ FILE REFERENCE: 062-001
/ CURRENT FILING DATE: 2001-03-20/227,860
/ PRIOR FILING DATE: 2000-09-24
/ PRIOR APPLICATION NUMBER: 09/667,569
/ NUMBER OF SEQ ID NOS: 77
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 45
/ LENGTH: 242
/ TYPE: PRF
/ ORGANISM: Bacillus anthracis
US-09-813-453A-45

Query Match      45.3%; Score 39; DB 9; Length 262;
Best Local Similarity 42.3%; Pred. No. 53;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      1 WKKKPRNLEEDYL 14
      |||
Db      26 WKMTDTRKTEDEY 39

RESULT 13
US-09-886-734A-12
/ Sequence 12; Application US/0986734A

```



```

/ Patent No. US20020164732A1
/ GENERAL INFORMATION: Marie-Christine
/ APPLICANT: Chappin, Marie-Christine
/ APPLICANT: Chappin, Marie-Christine
/ APPLICANT: Erlich, S Dinko
/ APPLICANT: Gaubier, Michel
/ APPLICANT: Schouler, Catherine
/ TITLE OF INVENTION: Institut National de la Recherche Agronomique
/ TITLE OF INVENTION: Resistance Mechanisms to ic Type R/M
/ TITLE OF INVENTION: Bacteriophages of Lactic Acid Bacteria
/ CURRENT FILING DATE: US/09/486,734A
/ PRIOR FILING DATE: 1998-09-01
/ PRIOR APPLICATION NUMBER: PCT/FR99/01873
/ PRIOR FILING DATE: 1997-09-02
/ NUMBER OF SEQ ID NOS: 40
/ SOFTWARE: FastSeq For Windows Version 4.0
/ LENGTH: 404
/ TYPE: PRT
/ ORGANISM: Lactococcus lactis
US-09-486-734A-12

Query Match 44.2%; Score 38; DB 9; Length 404;
Best Local Similarity 44.2%; Pred. No. 1.2e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 WKXPRRLLEDDYL 15
DB 50 KLFPSNREDDYVA 63

```

```

Query Match 44.2%; Score 38; DB 10; Length 539;
Best Local Similarity 46.2%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 WKXPRRLLEDDYL 15
DB 22 WFKXPGRDITNEDYL 36

```

```

RESULT 15
US-09-753-008-7
/ Sequence 7, Application US/09753008
/ Patent No. US20020061520A1
/ GENERAL INFORMATION:
/ APPLICANT: Stefan Somo and Tohiko Mochizuki
/ TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE PKD2 GENE
/ TITLE OF INVENTION: AND US95 THEROP
/ NUMBER OF SEQUENCES: 2
/ CORRESPONDING ADDRESS:
/ ADDRESSER: AMSTER, ROHMSTEIN & EBERNSTEIN
/ STREET: 90 PARK AVENUE
/ CITY: NEW YORK
/ STATE: NEW YORK
/ COUNTRY: U.S.A.
/ ZIP: 10016
/ COMPUTER HARDWARE FORM:
/ MEDIUM TYPE: 3.5-INCH 1.44 MB STORAGE
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: ASCII
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/753,008
/ FILING DATE: MAY 23, 1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/385,752
/ FILING DATE: 1999-08-30
/ APPLICATION NUMBER: 08/651,999
/ FILING DATE: MAY 23, 1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: ELIZABETH A. BOGOSIAN
/ REFERENCE/DOCKET NUMBER: 96700/395
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 697-5995
/ TELEFAX: (212) 286-0854 cr 286-0082
/ TELE: TWX 710-581-4766
/ INFORMATION FOR SEQ ID NO: 7
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 568 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: <unknown>
/ DESCRIPTION: peptide
/ HYPOTHETICAL: NO
/ SEQUENCES DESCRIPTION: SEQ ID NO: 7
US-09-753-008-7

```

```

Query Match 44.2%; Score 38; DB 10; Length 968;
Best Local Similarity 46.2%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 WKXPRRLLEDD 13
DB 293 WKKPNSQNGTADN 305

```

Search completed: March 26, 2003, 17:04:14
Job time : 11.1163 secs

US-09-815-242-10497
ORGANISM: Enterococcus faecalis


```

Db      24 WKMKPRNLEDDNDYL 38

RESULT 2
ID 092450;          PRELIMINARY;      PRT;      338 AA.
AC 092450;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE Six transmembrane endochelital antigen of PABC.
OS Six transmembrane endochelital antigen of PABC.
OC Mus musculus (Mouse).
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RA Nagasaka T., Boulay G., Coupel S., Coulon F., Tesson L.,
RA Hebrant J.-M., Soullivou J.-P., Charreau B.,
RA L'Heritier, gene expression in endochelital cells during TNF-alpha-
RA induced apoptosis in the liver of transgenic mice.
RA Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF319659; AAC33866.1; -.
KW Transmembrane.
SQ
SEQUENCE 338 AA; 39918 MW; ED49086067A329 CRC64;

Query Match      78.5%; Score 67.5; DB 6; Length 339;
Seq. Local Similarity 46.7%; 1; Mismatches 0; Indels 1; Gaps 1;
Matches 13; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy      1 WKMKPRNLEDDNDYL 15
Db      14 WKMKPRNLEDDNDYL 27
|||||:|||||
14 WKMKPRNLEDDNDYL 27

RESULT 3
092457;          PRELIMINARY;      PRT;      339 AA.
AC 092457;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE STEAP OR 241007181X.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RA Nagasaka T., Boulay G., Coupel S., Coulon F., Tesson L.,
RA Hebrant J.-M., Soullivou J.-P., Charreau B.,
RA L'Heritier, gene expression in endochelital cells during TNF-alpha-
RA induced apoptosis in the liver of transgenic mice.
RA Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF319659; AAC33866.1; -.
KW Transmembrane.
SQ
SEQUENCE 339 AA; 39918 MW; ED49086067A329 CRC64;

Query Match      78.5%; Score 67.5; DB 6; Length 339;
Seq. Local Similarity 46.7%; 1; Mismatches 0; Indels 1; Gaps 1;
Matches 13; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy      1 WKMKPRNLEDDNDYL 15
Db      14 WKMKPRNLEDDNDYL 27
|||||:|||||
14 WKMKPRNLEDDNDYL 27

```

```

DR MCD; MG1:1917608; Steap.
SQ SEQUENCE 339 AA; 39264 MW; 3F7AB9C7520F0968 CRC64;

Query Match      73.3%; Score 63; DB 11; Length 339;
Seq. Local Similarity 71.4%; Pred. No. 0.009; 2; Indels 0; Gaps 0;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 WKMKPRNLEDDNDYL 14
Db      14 WKMKPRNLEDDNDYL 27
|||||:|||||
14 WKMKPRNLEDDNDYL 27

RESULT 4
ID 092422          PRELIMINARY;      PRT;      339 AA.
AC 092422;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Six transmembrane endochelital antigen of PABC.
OS Six transmembrane endochelital antigen of PABC.
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RA Nagasaka T., Boulay G., Coupel S., Coulon F., Tesson L.,
RA Hebrant J.-M., Soullivou J.-P., Charreau B.,
RA L'Heritier, gene expression in endochelital cells during TNF-alpha-
RA induced apoptosis in the liver of transgenic mice.
RA Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF319659; AAC33866.1; -.
KW Transmembrane.
SQ
SEQUENCE 339 AA; 39109 MW; 32AC29F2E333800 CRC64;

Query Match      73.3%; Score 63; DB 11; Length 339;
Seq. Local Similarity 71.4%; Pred. No. 0.009; 2; Indels 0; Gaps 0;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      1 WKMKPRNLEDDNDYL 14
Db      14 WKMKPRNLEDDNDYL 27
|||||:|||||
14 WKMKPRNLEDDNDYL 27

RESULT 5
ID 092439          PRELIMINARY;      PRT;      339 AA.
AC 092439;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Six-transmembrane epithelial antigen of the prostate.
OS Six-transmembrane epithelial antigen of the prostate.
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RA Nagasaka T., Boulay G., Coupel S., Coulon F., Tesson L.,
RA Hebrant J.-M., Soullivou J.-P., Charreau B.,
RA L'Heritier, gene expression in endochelital cells during TNF-alpha-
RA induced apoptosis in the liver of transgenic mice.
RA Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF319659; AAC33866.1; -.
KW Transmembrane.
SQ
SEQUENCE 339 AA; 39318 MW; 4B26A71F5598B4F CRC64;

```

```

Query Match      73.3% Score 63.5 DB 11; Length 339;
Best Local Similarity 71.4%; Pred. No. 0.009; 2; Mismatches 0; Gaps 0;
Matches 10; Conservative 2; Indels 0;

QY      1 WKPKPRNLEDDY 14
      |||:|:|:|:|
DB      14 WKPKRKNLEDDSY 27

RESULT 6
QY      1 WKPKPRNLEDDY 14
AC      09UDU5      PRELIMINARY; PRT; 237 AA.
AC      09UDU5:
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT      01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DN      Conserved hypothetical protein.
GN      SPAC2358.15C.
OS      Schizosaccharomyces pombe (Fission yeast).
OC      Schizosaccharomycetes; Ascomycota; Fungi; Eukaryota;
OC      Schizosaccharomycetales; Schizosaccharomycetaceae;
OC      Schizosaccharomyces.
OX      NCBI_TaxID=4896;
RN      1
RP      SEQUENCE FROM N.A.
RA      Barrett B.G., Rajadream N.A., McDougall R.C., McLean J., Harris D.;
RA      BMC Microbiol. 3(2003)136. EMBL/GenBank/DBJ databases.
DR      EMBL; AL133225; CAB61781.1;
DR      InterPro; IPR003827; DUF207.
DR      Pfam; PF02676; DUF207.1.
DR      Hypothetical protein.
SQ      SEQUENCE 237 AA; 26397 MW; 29A6CFDDC720997 CRC64;

Query Match      52.3%; Score 45.5; DB 3; Length 237;
Best Local Similarity 66.7%; Pred. No. 6.9;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY      2 WKPKPRNLEDDY 15
      |||:|:|:|:|
DB      220 WKPKRKNLEDDYL 234

RESULT 7
QY      08XD7      PRELIMINARY; PRT; 375 AA.
AC      08XD7:
DT      01-MAR-2002 (TREMBLrel. 20, Created)
DT      01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DN      Klnase D, prokaryotic RNA precursor.
OS      Escherichia coli O157:H7.
OC      Escherichia coli O157:H7.
OC      Escherichia coli O157:H7.
OC      Bacterioid; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
OX      NCBI_TaxID=8334;
RN      11
RP      SEQUENCE FROM N.A.
RA      Kunitz M., Boudreau J., BDL03 / ATCC 700927;
RA      MEDLINE=1156915; PubMed=11705531;
RA      Perna N.T., Plunkett G.II, Burland V., Mau B., Glander J.D.,
RA      Rose D.J., Mayhew G.P., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA      Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA      Grobeck E.J., Davis N.W., Lin A., Dimalanta E.T., Potamousis K.,
RA      Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA      Welch R.A., Biltner P.R.;
RA      BMC Microbiol. 3(2003)136. EMBL/GenBank/DBJ databases.
RA      Medline 145359-533(2001).
RN      12
RP      SEQUENCE FROM N.A.
RA      STRAIN=O157:H7 / RIMD 0509952;
RA      MEDLINE=21156231; PubMed=1156796;
RA      Hayashi T., Makino K., Ohnishi W., Kurokawa K., Ishii K., Yokoyama K.,

```

```

RA      Han C.-G., Ohtsubo E., Nakayama K., Mureta T., Tanaka W., Tobe T.,
RA      Iida T., Takaki H., Honda T., Saekawa C., Ogasawara N., Yasunaga T.,
RA      Kunitake S., Shiba T., Hoshino K., Nakagawa H., Escherichia coli
RA      O157:H7 and genomic comparison with a laboratory strain K-12."
RT      DNA Res. 8:11-22(2001).
DR      EMBL; AE005403; AAC6793.1;
DR      EMBL; AP002558; BAB35936.1;
DR      InterPro; IPR002562; 3.5.exonuclease.
DR      Pfam; PF01812; HRC.
DR      Pfam; PF05570; HRC.exonuclease; 1.
DR      SMART; SM00341; HRC.1.
KW      Complete proteome.
SQ      SEQUENCE 375 AA; 42748 MW; 0D0F7D2982602BAF CRC64;

Query Match      52.3%; Score 45; DB 16; Length 375;
Best Local Similarity 63.6%; Pred. No. 14;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 WKPKPRNLE 11
      |||:|:|:|:|
DB      342 WKPKRKNLE 352

RESULT 8
QY      09LGG0      PRELIMINARY; PRT; 2058 AA.
AC      09LGG0:
DT      01-OCT-2000 (TREMBLrel. 15, Created)
DT      01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT      01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DN      P0406H10.14 protein.
OS      Oryza sativa (Rice).
OC      Oryza sativa (Rice).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Eukaryotes; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC      Ehrhartoideae; Oryzoae; Oryza.
OX      NCBI_TaxID=4530;
RN      11
RP      SEQUENCE FROM N.A.
RA      Sasaki T., Marumoto T., Yamamoto K.;
RA      Sasaki T., Marumoto T., Yamamoto T.;
RA      BMC Microbiol. 3(2003)136. EMBL/GenBank/DBJ databases.
RA      Sasaki T., Marumoto T., Yamamoto T.;
RT      "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT      clone:P0406H10."
RL      Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RA      EMBL; AP002524; BAB07964.1;
SQ      SEQUENCE 2058 AA; 22637 MW; 8A0B6562B8BD69 CRC64;

Query Match      52.3%; Score 45; DB 10; Length 2058;
Best Local Similarity 56.3%; Pred. No. 83;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      3 WKPKPRNLEDDY 14
      |||:|:|:|:|
DB      496 WKPKRKNLEDDY 507

RESULT 9
QY      09654      PRELIMINARY; PRT; 317 AA.
AC      09654:
DT      01-MAY-1999 (TREMBLrel. 10, Created)
DT      01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT      01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DN      Klnase C receptor homolog.
OS      Trypanosoma vivax.
OC      Trypanosoma vivax.
OC      Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX      NCBI_TaxID=5699;
RN      11
RP      SEQUENCE FROM N.A.
RA      STRAIN=JLDRLT.2;

```

[illegible]

ID	076517	PRLIMINARY;	PRT.	659 AA.
AC	01-NOV-1998	(TEMBLrel. 08, Created)		
DT	01-NOV-1998	(TEMBLrel. 08, Last sequence update)		
DR	01-DEC-2001	(TEMBLrel. 19, Last sequence update)		
DL	01-DEC-2001	(TrEMBLrel. 19, Last sequence update)		
DS	Trypanosoma brucei rhodesiane protein p67.			
OS	Trypanosoma brucei rhodesiane			
OC	Eukaryotes; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.			
NCBI	_taxid=31286;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RS	NC_009351.14; PubMed:10029105.			
RT	Keller R, Keller M, Alexander D L, Cowan C, Balher A E., Bangs J D.;			
RX	"Molecular cloning of p67, a lysosomal membrane glycoprotein from			
RL	Trypanosoma brucei."			
RM	Mol. Biochem. Parasitol. 98:17-28(1999).			
DR	EMBL; AF074867; AAC28456.1; "			
DN	InterPro; IPROU1064; Cytochrome b.			
DR	PROSITE; PS00225; CRYSTALLIN_BETANOMA; UNKNOWN_1			
DS	SEQUENCE 659 AA; 7270 MW; 483655C6105933 CRC64;			
OY	3 WKKRNLLEDYL 15	51.2%; Score 44; DB 5; Length 659;		
Dy		Best local similarity 61.5%; Pred.No. 37;		
	635 WKFRNRTEDELL 647	Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;		
RESULT 12				
ID	091UDS	PRELIMINARY;	PRT.	773 AA.
AC	09JUN05	(Created)		
DT	01-DEC-2001	(TEMBLrel. 19, Last sequence update)		
DR	01-DEC-2001	(TEMBLrel. 19, Last sequence update)		
DL	01-JUN-2002	(TRERELrel. 21, Last annotation update)		
DN	TOP protease.			
DS	TOP.			
OS	Rhizobium meliloti (Sinorhizobium meliloti).			
OC	Plantae; Proteobacteria; alpha subdivision; Rhizobiaceae group;			
NCBI	_taxid=5082;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RS	TRANSPOSOM-MERCURY RESISTANCE TRANSPOSON TNS718;			
RT	Schmeilner S., Keller M., Drosge M., Lanka E., Puchler A.,			
RX	Selitschnka W.;			
RL	The genetic organization and evolution of the broad-host-range			
RM	transposon TNS718 isolated from a microbial population			
DR	residing in the rhizosphere of alfalfa.";			
DN	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AJ304453; CAZ79182.1; "			
DN	InterPro; IPR002936; DNAPrim_toprim.			
DR	InterPro; IPR000380; Prok_Tprimase.			
DN	Ffam; PF01131; Topoisom_DacI.1.			
DR	Ffam; PF01131; Topoisom_DacI.1.			
DS	Transposon; TIR01051; topn_dact1.1.			
NCBI	_taxid=5082;			
SO	SEQUENCE 773 AA; 86188 MW; DBS3FPA6F20E68 CRC64;			
OY	Query Match	51.3%; Score 44; DB 2; Length 773;		
Dy	Best local similarity 53.2%; Pred.No. 44;			
	Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;			
RESULT 13				
ID	99 WHKESLSLDGYL 113	1 WKKRNLLEDYL 15		
		635 WKFRNRTEDELL 647		

FT SIGNAL 1 19 BY SIMILARITY.
 FT PROEP 20 107 ACTIVATION PEPTIDE (BY SIMILARITY).
 FT CHAIN 108 679 92 KDA TYPE IV COLLAGENASE.
 FT DOMAIN 109 105 FIBRINOLYTIC REGION.
 FT DOMAIN 298 105 FIBRONECTIN TYPE-II.
 FT DOMAIN 282 339 FIBRONECTIN TYPE-II.
 FT DOMAIN 340 397 FIBRONECTIN TYPE-II.
 FT DOMAIN 486 679 HEMOPEXIN-LIKE.
 FT METAL 401 401 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT SITE 402 402 BY SIMILARITY.
 FT METAL 403 405 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 431 431 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 491 676 BY SIMILARITY.
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 679 AA; 75561 MW; 573CDD99D8F85524 CRC64;

Query Match 50.04; Score 43; DB 13; Length 679;

Best Local Similarity 40.04; Fred. No. 57;

Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 WKKKPRRLREDDYL 15

Db 655 WRMTPEKQVDVGIV 669

Search completed: March 26, 2003, 16:49:42

Job time : 20.6377 secs

[illegible]

```

CC CC -----
CC CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation at
CC CC the European Bioinformatics Institute. These two institutions have agreed that
CC CC their respective contributions will remain in the public domain in no way that
CC CC could conflict with this statement; it is not removed. Usage by and for commercial
CC CC entities requires a license agreement. See http://www.ebi.ac.uk/annotation/
CC CC or send an email to license@ebi.ac.uk.
CC CC -----
DR DR EMBL; AB041768; BAB16722.1; -
DR DR EMBL; Z59168; CAB16301.1; -
DR DR EMBL; A5022743; RAB67247.1; --
DR DR InterPro: IPR000233; HSF DNA bind.
DR DR InterPro: IPR002341; HSF_ETS-
DR DR InterPro: IPR001789; Response-reg.
DR DR Pfam; PF00072; response_reg_1.
DR DR Pfam; PF00447; HSF-DNA-Bind; 1.
DR DR PRINTS; PR00056; HSFDOMAIN-reg_1.
DR DR PRODOM; PD001928; HSF-DNA_Bind; 1.
DR DR SMART; SMO0445; HSF_I.
DR DR SMART; SMO0448; REC; 1.
DR DR PROSITE; PS00434; HSF_DOMAIN; 1.
DR DR PROSITE; PSS0110; RESPONSE_REGULATOR; 1.
DR DR Transcription regulation; Sensory transduction; Nuclear protein;
DR DR DNA-binding; Phosphorylation BY SIMILARITY.
DR DR MOD_RES 369 481
DR DR MOD_RES 418 418 PHOSPHORYLATION (BY SIMILARITY).
DR DR SEQUENCE 539 AA; 600x4 MW; F384945C5B2871 CAC64;
SQ
Db Db Query Match 50.0%; Score 43; DB 1; Length 539;
Best Local Similarity 60.0%; Pred.No.10;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0.
Cy Cy 1 WKPKRRLVDEDEL 15
| : | | | | |
| : | | | | |
364 WKRPRLVVEDEL 378
-----
RESULT 5
ID ACNU BACSU STANDARD; PRT; 441 AA.
ODP2_BACSU
ACNU BACSU STANDARD; PRT; 441 AA.
AD P1863
01-PAY-1991 (Rel. 18, Created)
01-PAY-1991 (Rel. 18, Last annotation update)
01-JUN-2002 (Rel. 41, Last annotation update)
Dihydrolypidamide acetyltransferase component of pyruvate dehydrogenase
complex [EC 2.3.1.12] (E2) (S complex, 48 kDa subunit).
GN PHIC OR ACSC.
GN Bacillus subtilis
GN Bacillus subtilis
GN NCBI_Taxid=1423; Bacillales; Bacillaceae; Bacillus.
[1]
RP RP SEQUENCE FROM N.A.
RP STRAIN=168;
RP MEDLINE=90166558, PubMed=1597575;
RA Hemleir H., Palva A., Paulin L., Arvidson S., Palva T.;
RA Hemleir H., Palva A., Paulin L., Arvidson S., Palva T.;
RP Identity to pyruvate dehydrogenase".
RP J. Bacteriol. 172:5052-5063(1990)."
[2]
RP RP SEQUENCE FROM N.A.
RP STRAIN=168;
RP MEDLINE=97124187, PubMed=8869500;
RA Kinner S., Calabro R., Erdtman (de)g), Petzold E.;
RA Kinner S., Calabro R., Erdtman (de)g), Petzold E.;
RP Subtilisin 168 chromosome; sequencing of a 27 kb segment and
RP identification of several genes in the area.";
RL Microbiology 142:3033-3037(1996).
[3]
RP RP SEQUENCE FROM N.A.
```

CC STRAIN=168;
 RA Caldwell R.M., Petrait E.;
 RT Sequence analysis of the mobA-empS region of the *Bacillus subtilis*
 CC Submitted (Jul-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall
 CC conversion of pyruvate to acetyl-CoA and CO(2). It contains
 CC multiple copies of three enzymatic components: pyruvate
 CC dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and
 CC lipoamide dehydrogenase (E3). COMPLEX POSSESSES ALSO BRANCHED-CHAIN
 CC 2-OXOACID DEHYDROGENASE (BCHH) ACTIVITY
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + dihydrolipoamide = CoA + S-
 CC acetylhydrolipoamide.
 CC -1- COFACTOR: THE E2 COMPONENT CONTAINS ONE COVALENTLY-BOUND LIPOYL
 CC COFACTOR.
 CC -1- SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL
 CC SYMMETRY. BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.
 CC -1- SIMILARITY: CONTAINS LIPOYL-BINDING DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC or send a request to license@db-eb.ch (see <http://www.lsb-eb.ch/announce/>)
 CC -----
 CC EMBL: M57435; AAA62683.1; -;
 CC EMBL: AF012285; AAC24934.1; -;
 CC EMBL: Z59111; CAB1333.1; -;
 CC FPI: D36718; D36718;
 CC DR FPI: D36718; D36718;
 CC DR Subtilisin-B01209; ndhc
 CC InterPro: IPR001078; 2Oxoacid dh.
 CC InterPro: IPR000089; Biotin_lipoyl.
 CC InterPro: IPR004167; E3 binding.
 CC InterPro: IPR003161; Lipoyl.
 CC Pfam: PF00198; 2-oxoacid_dh; 1.
 CC Pfam: PF00364; biotin_lipoyl; 1.
 CC ProDom: PD01115; 2Oxoacid dh. 1.
 CC PROSITE: PS00189; LIPOYL; 1.
 CC PROSITE: PS00189; Transferase; Acyltransferase; Lipoyl; Complete proteome.
 CC INIT MET 0 0
 CC FT DOMAIN 1 85 LIPOYL BINDING.
 CC FT DOMAIN 139 181 E1/E3 BINDING.
 CC FT BINDING 412 412 LIPOYL.
 CC FT BINDING 412 412 LIPOYL.
 CC SEQUENCE 441 AA; 47407 MW; 684401879D134DF CRC64;
 SQ
 Query Match 48.8%; Score 42; DB 1; Length 441;
 Best Local Similarity 46.7%; Pred. No. 12;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 1 WKPKRNLEEDVYL 15
 Db 21 WKPKRNLEEDVYL 35
 RESULT 6
 ID ACUA_TRIRE STANDARD; PRT; 847 AA.
 AC Q99242; 1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Alpha-glucuronidase precursor (EC 3.2.1.139) (Alpha-glucosiduronase)
 CC (GLAI)
 CC Trichoderma reesei (Hypocrea jecorina).
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 CC Hypocrea; Hypocreales; Hypocrea.

CC NCBI: TaxID=51453;
 RX
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RX MEDLINE=66414; Pubmed=654984;
 RX MEDLINE=9625277; Pubmed=654984;
 RT Margolis-Clark E., Satohelmo M., Sika-Aho M., Penttilae M.;
 RL Genet 172:171-172(1996).
 CC -1- FUNCTION: RELEASES 4-O-METHYLGALACTURONIC ACID FROM XYLAN.
 CC -1- CATALYTIC ACTIVITY: An alpha-D-glucuronoside + H(2)O = an alcohol
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO FAMILY 67 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC or send a request to license@db-eb.ch (see <http://www.lsb-eb.ch/announce/>)
 CC -----
 CC EMBL: Z68706; CAA92949.1; -;
 CC InterPro: IPR005154; Glyco_hydro. 67.
 CC Pfam: PF03648; Glyco_hydro. 67; 1.
 CC Xylan degradation; Hydrolase; Glycosidase; Glycoprotein; Signal.
 CC KJ
 CC FT CHAIN 20 847
 CC FT CARBOHYD 52 52 N-LINKED (GLCNAC...) (POTENTIAL)
 CC FT CARBOHYD 238 238 N-LINKED (GLCNAC...) (POTENTIAL)
 CC FT CARBOHYD 321 321 N-LINKED (GLCNAC...) (POTENTIAL)
 CC FT CARBOHYD 352 352 N-LINKED (GLCNAC...) (POTENTIAL)
 CC FT CARBOHYD 366 366 N-LINKED (GLCNAC...) (POTENTIAL)
 CC FT CARBOHYD 586 586 N-LINKED (GLCNAC...) (POTENTIAL)
 CC FT CARBOHYD 740 740 N-LINKED (GLCNAC...) (POTENTIAL)
 CC FT CARBOHYD 767 767 N-LINKED (GLCNAC...) (POTENTIAL)
 CC SEQUENCE 847 AA; 93424 MW; 214AAE8732ED84 CRC64;
 SQ
 Query Match 48.8%; Score 42; DB 1; Length 847;
 Best Local Similarity 53.8%; Pred. No. 26;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 2 KMPKRNLEEDVYL 14
 Db 108 KMPKRNLEEDVYL 120
 RESULT 7
 ID ACUA_TRIRE STANDARD; PRT; 1018 AA.
 AC Q99242; 1998 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Serine/threonine protein kinase 31 (EC 2.7.1.37).
 CC (SKK1)
 CC Mus musculus (mouse)
 CC Chordata; Carnifera; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI: TaxID=10090.
 RX
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21115748; Pubmed=11279525;
 RX Wang R.-J., McGilley R.-J., Tang F., Page D. C.;
 RL Nat. Genet. 27:442-446(2001)
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- TISSUE SPECIFICITY: Testis-specific. Expressed only in male germ
 CC cells.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- SIMILARITY: CONTAINS 1 TUDOR DOMAIN.
 CC -----

CC	This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL collaboration.
CC	The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way misrepresented and any reference to it must acknowledge the source and copyright holder.
CC	EMBL; AF285580; AAK31959.1 / -
DR	MGI; MGI:1924735; SLKX1.
DR	InterPro; IPR000719; Elk_kinase.
DR	InterPro; IPR001097; Metacore_Tudor.
DR	InterPro; IPR002280; Serine_kinase.
DR	InterPro; IPR002899; Tudor.
DR	Pfam; PF00669; pkinase; 1.
DR	Pfam; PF00567; TUDOR; 1.
DR	ProDom; PD000001; Elk_kinase; 1.
DR	SMART; SMART0220; S_TGc; 1.
DR	SMART; SMART0333; TUDOR_1.
DR	PROSITE; PS0004; TUDOR_KINASE_DOM; 1.
DR	Transferrase; Serine/threonine-protein kinase; ATP-binding; Coiled coil.
KV	Coiled coil. 78 137 TUDOR.
FW	DOMAIN FT 105 245 THASE-COIL.
FT	DOMAIN FT 298 358 COILED COIL (POTENTIAL).
FT	DOMAIN FT 1 1018 PROTEIN KINASE_SITRY.
FT	BINDING FT 717 738 ATP (BY SIMILARITY).
FT	BINDING FT 738 738 ATP (BY SIMILARITY).
SQ	SEQUENCE 1018 AA; 115027 MW; 1206284ED4E90816 CRC64;
Qy	Query Match 48.8%; Score 42; DB 1; Length 1018; Best Local Similarity 53.8%; Pred. No. 31/ Matches 1; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Dh	646 WKVKSKNSLESD 658 1 WKKGRNTHLED 13
RESULT 8	
ID	ST31 HUMAN STANDARD; PRT; 1019 AA.
GN	ST31_HUMAN
OS	Homo sapiens (Human)
OC	Eukaryota; Chordata; Carnivora; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
NCBI_TaxID=9606;	[1]
RN	SEQUENCE FROM N.A.
RC	TISSUE=Tes1a;
RX	MEDLINE=21175748; PubMed=1279525;
RA	Wang F.-J., McCreary V.R., Yang F.; Page D.C.; Nat Biotechnol. 2003; 21(12):1201-1206. [2] Gene. 27(42):24612001.
RN	SEQUENCE OF 27-42-1019 FROM N.A.
RC	TISSUE=Tes1a;
RT	Zhou Z.M.;
RT	"Cloning of a new protein kinase gene related to human testis development." (May 2001) to the EMBL/GeneDB/DDI database.
CC	-1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC	-1- TISSUE SPECIFICITY: Testis-specific.
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC	-1- SIMILARITY: CONTAINS 1 TUDOR DOMAIN.

[illegible]

PT PTOEALPESHSHSPQOC (IN ISOFORM A).
 FT VARSPPLIC 77 113
 FT NGTSLSLCAVCSRPFGITLWLNKSNFPHLPGLM ->
 FT SNAENSLNAPHPASALQCSYALGRLSLTPALAAP (IN
 FT ISOFORM B)
 FT VARSPPLIC 114 197
 FT MESTLSP (IN ISOFORM B).
 SO SEQUENCE 197 AA; 21698 MW; P1B3E26A82B395EL CRC64;
 Query Match 46.5%; Score 40; DB 1; Length 197;
 Best Local Similarity 46.7%; Pred. No. 11;
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 QY 1 WKKKRPRNLEED 15
 Db 167 WYSPRNLQDZEL 181
 RESULT 14
 ID STR_ARCPV STANDARD; PRT; 549 AA.
 DC 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ARGINYL-tRNA synthetase [EC 6.1.1.19] (Arginine--tRNA ligase) (ArgRS).
 GN AAGS OR AF0894.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 NC Archaeoglobus fulgidus.
 RX MEDLINE:98049343; PubMed:938975;
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RA Klenk H.-P., Clayton R.A., Tomb J.-P., White O., Nelson K.E.,
 RA Richardson A., Dodson J., McQuinn K., Eickbush S.H.,
 RA Fleischmann R.D., Chacko J., Smith R.A., Peterson D.E., Karpas NC,
 RA Peterson S.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Ralch C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne T.J., Weidman J.F., McDonald L., Utterback T.,
 RA Sadow P.M., Olsen G.J., Fraser C.M., Smith R.A., Karp S.A.,
 RA Venter J.C., Olsen G.J., Fraser C.M., Smith R.A., Karp S.A.,
 RA "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 CC -1- CATALYTIC ACTIVITY: ATP + L-arginine + tRNA(Arg) = AMP +
 CC dihydrophosphate + L-arginyl-tRNA(Arg).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC EMBL: AF0894.1; AAB90346.1; -.
 DR TIGR: AF0894.1;
 DR InterPro: IPR001278; Arg_rRNA-synt_1c.
 DR InterPro: IPR005148; N.
 DR InterPro: IPR001812; tRNA-synt_1.
 DR Pfam: PF03485; tRNA-synt_1d; 1.
 DR Pfam: PF03485; tRNA-synt_1d; 1.
 DR PRINTS: PR01038; TRNASYTHAG.
 DR TIGRFAMs: TIGR00455; args_1.
 DR PROSITE: PS00178; AA_tRNA_LIGASE_I; FALSE_NEG.
 DR Anticopy1-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome. 123
 FT SITE 113 "HIGH" REGION.

SQ SEQUENCE 549 AA; 62861 MW; 39DF41CC0B9AD210 CRC64;
 Query Match 46.5%; Score 40; DB 1; Length 549;
 Best Local Similarity 46.2%; Pred. No. 3;
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 1 WKKKRPRNLEED 13
 Db 377 WKLSRMEDE 389
 RESULT 15
 ID GPTI_CANAL STANDARD; PRT; 553 AA.
 AC 074248;
 DT 15-JUL-1998 (Rel. 38, Created)
 DT 15-JUN-1998 (Rel. 38, Last sequence update)
 DT 15-JUN-1998 (Rel. 38, Last annotation update)
 DE Putative polyamine transporter.
 GN Candida albicans (Yeast).
 OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetes; mtosporic Saccharomycetales; Candida.
 NC Candida albicans.
 RX NCBI_TaxID=5476;
 RN 1)
 RP SEQUENCE FROM N.A.
 RA McKenna M.D., Gorman J.A., Buckley H.R.,
 RA Buchanan M.P., The GPT1 gene of *Candida albicans*, encoding
 RT a putative polyamine transporter.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC EMBL: F0690293.1; AF035338.1; AF035338.1;
 DR InterPro: IPR004840; ABC_permease.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00324; aa_permeases; 1.
 DR PROSITE: PS00218; AMINO ACID PERMEASE_1; FALSE_NEG.
 KW Transport; Transmembrane.
 FT TRANSMEM 43
 FT TRANSMEM 133
 FT TRANSMEM 153
 FT TRANSMEM 163
 FT TRANSMEM 190
 FT TRANSMEM 240
 FT TRANSMEM 260
 FT TRANSMEM 280
 FT TRANSMEM 300
 FT TRANSMEM 333
 FT TRANSMEM 353
 FT TRANSMEM 397
 FT TRANSMEM 417
 FT TRANSMEM 485
 FT TRANSMEM 505
 SO SEQUENCE 553 AA; 197DCA77C6SFBC9C CRC64;
 Query Match 46.5%; Score 40; DB 1; Length 553;
 Best Local Similarity 46.2%; Pred. No. 35;
 Matches 6; Conservative 3; Mismatches 0; Gaps 0;
 QY 1 WKKKRPRNLEED 13
 Db 508 WYGRPRNLEED 520
 Search completed: March 26, 2003, 16:45:03
 Job time : 5.36047 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Comogen Ltd.

OW protein - protein search, using sw model

Run on: March 26, 2003, 16:44:15 ; Search time 8.54651 seconds

(without alignments)
168.726 Million cell updates/sec

Title: US-10-010-667a-2_COPY_14_28

Sequence: 1 WKPKPRRLSEDDYL 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	45.5	237	2 T50202	conserved hypothetical protein SPAC2588.15c [Imported] - fission yeast (Schizosaccharomyces pombe)
2	45.5	237	1 R5C002	ribonuclease III (RNase D, processes
3	45.5	237	2 A99943	probable heat shock protein 70 kDa (Hsp70)
4	45.5	237	2 B85791	hypothetical protein
5	45.5	237	2 T02539	hypothetical protein
6	45.5	237	2 T15371	probable heat shock protein 70 kDa (Hsp70)
7	45.5	237	2 T15371	hypothetical protein
8	45.5	237	2 T15371	hypothetical protein
9	45.5	237	2 T15371	hypothetical protein
10	45.5	237	2 T15371	hypothetical protein
11	45.5	237	2 T15371	hypothetical protein
12	45.5	237	2 T15371	hypothetical protein
13	45.5	237	2 T15371	hypothetical protein
14	45.5	237	2 T15371	hypothetical protein
15	45.5	237	2 T15371	hypothetical protein
16	45.5	237	2 T15371	hypothetical protein
17	45.5	237	2 T15371	hypothetical protein
18	45.5	237	2 T15371	hypothetical protein
19	45.5	237	2 T15371	hypothetical protein
20	45.5	237	2 T15371	hypothetical protein
21	45.5	237	2 T15371	hypothetical protein
22	45.5	237	2 T15371	hypothetical protein
23	45.5	237	2 T15371	hypothetical protein
24	45.5	237	2 T15371	hypothetical protein
25	45.5	237	2 T15371	hypothetical protein
26	45.5	237	2 T15371	hypothetical protein
27	45.5	237	2 T15371	hypothetical protein
28	45.5	237	2 T15371	hypothetical protein
29	45.5	237	2 T15371	hypothetical protein

30	40	46.5	549	1 F63561	arginyl-tRNA synth
31	40	46.5	884	2 H83122	hypothetical prote
32	40	46.5	1268	2 T10555	hypothetical prote
33	40	46.5	1268	2 T10555	hypothetical prote
34	39	45.3	90	2 H90655	probable Ogt fam11
35	39	45.3	90	2 H90655	probable activator
36	39	45.3	371	2 C64499	isocitrate dehydro
37	39	45.3	426	2 B83981	pyruvate dehydroge
38	39	45.3	426	2 B83981	protein t26k.2 (lm
39	39	45.3	449	2 G89841	hypothetical prote
40	39	45.3	449	2 G89841	hypothetical prote
41	39	45.3	512	2 T23035	hypothetical prote
42	39	45.3	652	2 B84568	probable calmodul
43	39	45.3	713	2 JCS670	poly(beta-D-mannu
44	39	45.3	976	2 A97104	Zn-dependent metal
45	39	45.3	1019	2 T00117	dve protein - fru1

ALIGNMENTS

RESULT 1

T50202

conserved hypothetical protein SPAC2588.15c [Imported] - fission yeast (Schizosaccharomyces pombe)

C/Date: 09-Jun-2000 #sequence, revision 09-Jun-2000 #ext. change 09-Jun-2000

R/Accession: T50202 #seq. rev. 09-Jun-2000 #ext. change 09-Jun-2000

R/Accession: T50202 #seq. rev. 09-Jun-2000 #ext. change 09-Jun-2000

submitted to the EMBL Data Library, November 1999

A/Reference number: Z25045

A/Accession: T50202

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Accession: T50202

A/Accession: T50202

A/Accession: T50202

A/Accession: T50202

A/Accession: T50202

A/Accession: T50202

A/Accession: T50202

A/Accession: T50202

A/Accession: T50202

A/Accession: T50202

A/Accession: T50202

A/Accession: T50202

A/Accession: T50202

A/Accession: T50202

A/Accession: T50202

A/Accession: T50202

A/Accession: T50202

A/Accession: T50202


```

C:/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:/Accession: T15371
R:/Johnson, D.
submitted to the EMBL Data Library, May 1996
A:/Description: the sequence of C. elegans cosmid C01F1.
A:/Accession: T15371
A:/Accession number: 210538
A:/Status: preliminary; translated from GB/EMBL/DBJ
A:/Molecule type: DNA
A:/Residues: 1,312 <0>H>
A:/Cross-references: EMBL:U58761; NID:G130391; PID:G1330396; PIDN:AA80076.1; GSPDB:GN0C
A:/Experimental source: strain Bristol N2; Clone C01F1
A:/Genetic CSP: C01F1.2
A:/Map position: 2
A:/Introns: 36/2; 271/1

Query Match          50.0%; Score 43; DB 2; Length 312;
Best Local Similarity 63.6%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 5 PRINTEDDY 15
|||:||||:
Db 250 PRYKQEDDY 260

RESULT 7
T33150
C:/Species: Schizosaccharomyces pombe
C:/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:/Accession: T33150
R:/Olliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:/Accession: T33150
A:/Accession number: Z21748
A:/Molecule type: DNA
A:/Residues: 1,539 <0>LI>
A:/Cross-references: EMBL:Z99168; PIDN:CA61301.1; GSPDB:GN00066; SPDB:SPAC8C9.14
A:/Experimental source: strain 972H; cosmid c8C9
C:/Genetics:
A:/Gene: SPDB:SPAC8C9.14
A:/Map position: 1
A:/Introns: 10/3; 40/3; 67/2; 86/3

Oy 1 WAKGRRNLEDDY 15
Db 364 WGRNRLVWDEDL 378

RESULT 8
AH1754
C:/Species: Listeria innocua
C:/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:/Accession: AH1754
R:/Gisener, P.; Frangoul, L.; Buchliesser, C.; Amand, A.; Baquero, F.; Berche, P.; Bloekerker,
A.; Domerg, L.M.; Kerez, U.
Science 294, 849-852, 2001
A:/Authors: Kerez, U.; Kuhn, M.; Kuntz, F.; Kuzpjak, G.; Madueno, E.; Matlounan, A.; M
K.; C.; Schlueter, J.; Simoes, J.; Tietzer, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland
A:/Reference number: AB1077; NID:21537279; PMID:11679669
A:/Accession: AH1754
A:/Status: preliminary
A:/Molecule type: DNA
A:/Residues: 1-124 <GLA>

```

A:crosses-references:	GB:AL592022; PTDN:OAC97808.1; PID:g16415103; GSPDB:GN00178
C:Species:	Arabidopsis thaliana (mouse-ear crests)
A:Experimental source:	strain C1p1262
C:Genetic(s):	
A:Gene:	lin2581
Query Match	49.8%; Score 42; DB 2; Length 124;
Best Local Similarity	46.7%; Pred. No. 8.6;
Matches	7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Oy	1 WNKPRRLIEDYLL 15
Db	15 WKRRKXILRDEYL 29
RESULT 9	
T49020	
Hypothetical protein F3C22_70 - Arabidopsis thaliana (fragment)	
C:Species:	Arabidopsis thaliana (mouse-ear crests)
C:Date:	02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
R:Accession:	F34920
A:Submitted to the Protein Sequence Database,	April 2000
A:Reference number:	225013
A:Accession:	T49020
A:Status:	Preliminary
A:Molecule type:	DNA
A:Cross-references:	EMBL:ALJ5312; GSPDB:GN00061; ANSP:F3C22_70
C:Genetic(s):	
A:Experimental source:	cultivar Columbia; BAC clone F3C22
A:Map position:	3
A:Interons:	61/1; 252/3; 281/3
Query Match	49.8%; Score 42; DB 2; Length 384;
Best Local Similarity	42.3%; Pred. No. 2;
Matches	6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
Oy	1 WNKPRRLIEDDY 14
Db	46 WKLEPFLFSDDY 59
RESULT 10	
D56532	
Hypothetical protein T1217.9 [imported] - Arabidopsis thaliana	
C:Species:	Arabidopsis thaliana (mouse-ear crests)
C:Date:	02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 11-Mar-2001
R:Accession:	D56532
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federmyer, N.A.; Kaul, S.; White, O.; Alonso, Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creamy, T.H.; Dewar, K.; Nature 408, 816-820, 2000	
A:Authors:	Junker, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, S.; Kim, C.; Rizzo, M.; Rooney, T.; Rowley, D.; Sekano, H.
A:Authors:	Salzberg, S.L.; Schwartz, J.C.; Shim, P.; Southwick, A.M.; Sun, H.; Talton, I.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.	
A:Title:	Sequence and analysis of chromosome of the plant Arabidopsis.
A:Accession:	D56532
A:Accession:	AB0141; U00212016713; FMD1110V12
A:Molecule type:	DNA
A>Status:	Preliminary
A:Gene:	T1217.9
A:Map position:	1
Query Match	48.8%; Score 42; DB 2; Length 416;
Best Local Similarity	50.0%; Pred. No. 3;
Matches	7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Thu Mar 27 09:53:34 2003

```
QY      3 MKPRNLEEDDYL 15
          ::|||:|||||
Db     356 LRPLRTVETDDYL 366
```

RESULT 15

hypothetical protein AGR_C_235 [Imputed] Agrobacterium tumefaciens (strain C58, Ceratophyllaceae)
C.Species: Agrobacterium tumefaciens
C.Date: 30-Sep-2001 [sequence_revision 10-Sep-2001 |text change 11-Jan-2002
C.Accession: D97376
R.Louder, B. Hinkle, G. Gattung, S. Miller, N. Blanchard, M. Qurollo, B. Goldman,
A. Lu, F. Molian, C. Allinger, M. Doughty, D. Scott, C. Lappas, C. Marketz, B.
A. Patel, Genome Sequencing Center, The Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A.Accession: D97376
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 11858 <RMS>
A.Cross-references: GB:AE007869; PID:AAK5965.1; PID:g1515024; GSPDB:GN00169
C.Genetic:
A.Gene: AGR_C_235
A.Map position: circular chromosome

Query Match:

Query Match	48.8%;	Score 42;	DB 2;	Length 1858;
Best Local Similarity	53.8%;	Pred. No. 1.5e+02;		
Matches	7;	Conservative	3;	Mismatches 3; Indels 0; Gaps 0;

QY 3 MKPRRLLEDDYL 15

Db 382 LRPLRTVETDDYL 394

Search completed: March 26, 2003, 16:50:36
Job time : 12.5465 secs



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 16:49:57 ; Search time 10.1163 Seconds
(without alignments)
87.073 Million cell updates/sec

Title: US-10-010-667a-21
Perfect score: 93

Sequence: 1 WIDIXQFWYTPPTF 15

Scoring table: BLASTNM62
Gapop 10.0, Gapext 0.5

Searched: 237916 seqs, 58723574 residues

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing filter 49 summaries

Database :

Published Applications AA:*
1: /cgn2_6/prodata/1/pubpaap/US08_NEW_PUB.pep:*
2: /cgn2_6/prodata/1/pubpaap/PCT_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubpaap/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/1/pubpaap/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaap/US07_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubpaap/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubpaap/US08_PUBCOMB.pep:*
8: /cgn2_6/prodata/1/pubpaap/US09_NEW_PUB.pep:*
9: /cgn2_6/prodata/1/pubpaap/US09_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubpaap/US10_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubpaap/US10_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubpaap/US10_PUBCOMB.pep:*
13: /cgn2_6/prodata/1/pubpaap/US10_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubpaap/US10_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score of the result being printed, and is derived by analysis of the score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	100.0	15	9 US-10-011-095-21	Sequence 21, Appl1
2	93	100.0	15	9 US-10-010-667a-21	Sequence 21, Appl1
3	93	100.0	267	10 US-09-147-835a-50	Sequence 50, Appl1
4	93	100.0	267	10 US-09-147-835a-51	Sequence 51, Appl1
5	93	100.0	339	9 US-09-802-520-11	Sequence 11, Appl1
6	93	100.0	339	9 US-09-802-520-11	Sequence 11, Appl1
7	93	100.0	339	9 US-09-895-793-079	Sequence 879, Appl1
8	93	100.0	339	9 US-09-895-814-079	Sequence 879, Appl1
9	93	100.0	339	9 US-10-011-095-2	Sequence 2, Appl1
10	93	100.0	339	9 US-10-010-667a-2	Sequence 2, Appl1
11	93	100.0	339	10 US-08-159-173-879	Sequence 879, Appl1
12	93	100.0	339	10 US-08-159-173-879	Sequence 879, Appl1
13	93	100.0	339	10 US-08-159-173-879	Sequence 879, Appl1
14	93	100.0	268	10 US-09-854-816-33	Sequence 33, Appl1
15	93	100.0	268	10 US-09-854-816-33	Sequence 33, Appl1
16	93	100.0	568	10 US-09-815-242-5492	Sequence 5492, Appl1
17	93	100.0	568	10 US-09-815-242-5492	Sequence 5492, Appl1
18	93	100.0	447	9 US-09-873-134-2	Sequence 47, Appl1
19	93	100.0	351	9 US-10-026-741-47	Sequence 47, Appl1

20	43	46.2	873	1 US-08-911-824-61	Sequence 61, Appl1
21	43	46.2	877	9 US-10-026-741-102	Sequence 102, Appl1
22	42	45.2	268	10 US-09-854-816-26	Sequence 26, Appl1
23	42	45.2	268	10 US-09-854-816-26	Sequence 26, Appl1
24	42	45.2	268	10 US-09-854-816-26	Sequence 26, Appl1
25	42	45.2	459	9 US-10-007-527a-21	Sequence 21, Appl1
26	42	45.2	619	10 US-09-891-609-4	Sequence 4, Appl1
27	42	45.2	646	10 US-09-891-609-2	Sequence 2, Appl1
28	42	45.2	847	10 US-09-176-742-2	Sequence 2, Appl1
29	41	44.1	243	10 US-09-854-816-73	Sequence 73, Appl1
30	41	44.1	268	10 US-09-854-816-73	Sequence 73, Appl1
31	41	44.1	268	10 US-09-854-816-73	Sequence 73, Appl1
32	41	44.1	268	10 US-09-854-816-73	Sequence 73, Appl1
33	41	44.1	268	10 US-09-854-816-73	Sequence 73, Appl1
34	41	44.1	268	10 US-09-854-816-73	Sequence 73, Appl1
35	41	44.1	268	10 US-09-854-816-73	Sequence 73, Appl1
36	41	44.1	268	10 US-09-854-816-73	Sequence 73, Appl1
37	41	44.1	268	10 US-09-854-816-73	Sequence 73, Appl1
38	41	44.1	268	10 US-09-854-816-73	Sequence 73, Appl1
39	41	44.1	268	10 US-09-854-816-73	Sequence 73, Appl1
40	41	44.1	351	9 US-09-886-150-46	Sequence 46, Appl1
41	41	44.1	351	9 US-09-886-150-46	Sequence 46, Appl1
42	41	44.1	351	9 US-09-886-150-46	Sequence 46, Appl1
43	41	44.1	351	9 US-09-886-150-46	Sequence 46, Appl1
44	41	44.1	625	9 US-10-032-162-17	Sequence 16, Appl1
45	41	44.1	655	10 US-09-725-735a-12	Sequence 12, Appl1

ALIGNMENTS

RESULT 1
US-10-011-095-21
Sequence 21, Application US/10011095
US-10-010-667a-21
Sequence 21, Application US/10010667a21
APPLICANT: Alcar, Daniel
APPLICANT: Hubert, Rene S.
APPLICANT: Leong, Kahan
APPLICANT: Raitano, Arthur B.
APPLICANT: Saitani, Douglas C.
APPLICANT: Mitchell, Steve (Chappell)
APPLICANT: Mitchell, Steve (Chappell)
FILE REFERENCE: S1116201610
CURRENT APPLICATION NUMBER: US/10/011.095
PRIOR FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: 09/323.873
PRIOR FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: 60/087.520
PRIOR FILING DATE: 1998-06-01
PRIOR APPLICATION NUMBER: 60/091.183
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 21
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
OTHER INFORMATION: STEAP-1 PEPTIDE
US-10-011-095-21

Query Match 100.0%; Score 93; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.5e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 WIDIXQFWYTPPTF 15
|||||
1 WIDIXQFWYTPPTF 15

RESULT 2
US-10-010-667a-21

```

/ Sequence 21, Application US/10010667a
/ Publication No. US20030055217A1
/ GENERAL INFORMATION:
/ APPLICANT: Mar, Daniel
/ APPLICANT: Mar, Daniel S.
/ APPLICANT: Raitano, Arthur B.
/ APPLICANT: Raitano, Arthur B.
/ APPLICANT: Saffran, Douglas C.
/ APPLICANT: Mitchell, Steve Chappell
/ TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
/ FILE REFERENCE: HVS-37CIP
/ CURRENT FILING DATE: 2001-12-06
/ PRIOR APPLICATION NUMBER: 09/223,873
/ PRIOR FILING DATE: 1999-06-01
/ PRIOR APPLICATION NUMBER: 60/087,520
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/091,183
/ PRIOR FILING DATE: 1998-06-30
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: PatSeq for Windows Version 4.0
/ SEQ ID NO 21
/ LENGTH: 15
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ ORGANISM INFORMATION: STAB-1 PEPTIDE
US-10-010-667a-21

Query Match          100.0%; Score 93; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 4,5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 WIDKQFWWTPTPF 15
Db      1 WIDKQFWWTPTPF 15

RESULT 3
/ Sequence 50, Application US/09747835A
/ Publication No. US2002018692A1
/ GENERAL INFORMATION:
/ APPLICANT: Yamazaki, Victoria
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Zhou, Ping
/ APPLICANT: Wang, Dunhui
/ APPLICANT: Wang, Jie
/ APPLICANT: Ren, Feiyun
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Dmanac, Radjoe T
/ TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
/ FILE REFERENCE: HVS-37CIP
/ CURRENT APPLICATION NUMBER: US/09/747,835A
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: US 09/653,450
/ PRIOR FILING DATE: 2000-12-04
/ PRIOR APPLICATION NUMBER: US 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: US 09/598,042
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: US 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: US 09/486,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 63
/ SOFTWARE: Patentin version 3.0
/ SEQ ID NO 50

```

```

/ LENGTH: 267
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-747-835a-50

Query Match          100.0%; Score 93; DB 10; Length 267;
Best Local Similarity 100.0%; Pred. No. 8,2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 WIDKQFWWTPTPF 15
Db      211 WIDKQFWWTPTPF 225

RESULT 4
/ Sequence 51, Application US/09747835A
/ Publication No. US2002018692A1
/ GENERAL INFORMATION:
/ APPLICANT: Yamazaki, Victoria
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Zhou, Ping
/ APPLICANT: Wang, Dunhui
/ APPLICANT: Wang, Jie
/ APPLICANT: Ren, Feiyun
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Dmanac, Radjoe T
/ TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
/ FILE REFERENCE: HVS-37CIP
/ CURRENT APPLICATION NUMBER: US/09/747,835A
/ PRIOR FILING DATE: 2002-05-09/729,739
/ PRIOR FILING DATE: 2000-12-04
/ PRIOR APPLICATION NUMBER: US 09/653,450
/ PRIOR FILING DATE: 2000-08-31
/ PRIOR APPLICATION NUMBER: US 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: US 09/598,042
/ PRIOR FILING DATE: 2000-06-01
/ PRIOR APPLICATION NUMBER: US 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: US 09/486,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 63
/ SOFTWARE: Patentin version 3.0
/ SEQ ID NO 51
/ LENGTH: 267
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-747-835a-51

Query Match          100.0%; Score 93; DB 10; Length 267;
Best Local Similarity 100.0%; Pred. No. 8,2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 WIDKQFWWTPTPF 15
Db      211 WIDKQFWWTPTPF 225

RESULT 5
/ Sequence 57, Application US/10012896
/ Publication No. US2002018625A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Dillon, David C.
/ APPLICANT: Mitchell, Jennifer L.
/ APPLICANT: Hatlocker, Susan L.
/ APPLICANT: Jiang, Yuguu
/ APPLICANT: Kalon, Michael D.
US-10-012-896-879

```



```

/ APPLICANT: Retter, Marc W.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darriek
/ APPLICANT: Wang, Aijun X.
/ APPLICANT: Skeiky, Yashir A.W.
/ APPLICANT: Hepler, William T.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Hural, John
/ APPLICANT: Houghton, Raymond L.
/ APPLICANT: Vinals de Bassols, Carlota
/ APPLICANT: Foy, Teresa
/ APPLICANT: Panger, Gary R.
/ APPLICANT: Mantanebe, Yoshitiro
/ APPLICANT: Meagher, Madeline Joy
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: COMPOSITIONS OF PROSTATE CANCER
/ FILE REFERENCE: 210121.427C27
/ CURRENT APPLICATION NUMBER: US/10/012.896
/ CURRENT FILING DATE: 2001-12-10
/ NUMBER OF SEQ ID NOS: 1011
/ SOFTWARE: FASTSEQ for Windows Version 3.0
/ SEQ ID NO 879
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-012-896-879

```

```

Query Match          100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 WIDKOPFWTPTPTF 15
Db 277 WIDKOPFWTPTPTF 291

```

```

RESULT 6
US-09-802-520-11
/ Sequence 11, Application US/09802520
/ Publication No. US2002018742A1
/ GENERAL INFORMATION:
/ APPLICANT: Lai, Preci
/ APPLICANT: Faris, Mary
/ APPLICANT: Chen, Hwei-Mei
/ APPLICANT: Ison, Craig H.
/ TITLE OF INVENTION: STRAP-RELATED PROTEIN
/ FILE REFERENCE: PC-0037 US
/ CURRENT APPLICATION NUMBER: US/09/802.520
/ CURRENT FILING DATE: 2001-03-09
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: PERL Program
/ SEQ ID NO 11
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Inocyte ID No. US2002018742A1 66572948
US-09-802-520-11

```

```

Query Match          100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 WIDKOPFWTPTPTF 15
Db 277 WIDKOPFWTPTPTF 291

```

```

RESULT 7
US-09-895-793-879
/ Sequence 879, Application US/09895793
/ Publication No. US20020192765A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Dillon, Devin C.
/ APPLICANT: Mitchell, Jennifer L.
/ APPLICANT: Harlocker, Susan L.
/ APPLICANT: Ulang, Yugu
/ APPLICANT: Ralston, Michael D.
/ APPLICANT: Houghton, Robert A.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darriek
/ APPLICANT: Wang, Aijun X.
/ APPLICANT: Skeiky, Yashir A.W.
/ APPLICANT: Hepler, William T.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Hural, John
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Vinals de Bassols, Carlota
/ APPLICANT: Foy, Teresa
/ APPLICANT: Panger, Gary R.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
/ FILE REFERENCE: 210121.534C2
/ CURRENT APPLICATION NUMBER: US/09/895.793
/ CURRENT FILING DATE: 2001-06-29
/ NUMBER OF SEQ ID NOS: 982
/ SOFTWARE: FASTSEQ for Windows Version 3.0
/ SEQ ID NO 935
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-895-793-879

```

```

Query Match          100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 WIDKOPFWTPTPTF 15
Db 277 WIDKOPFWTPTPTF 291

```

```

RESULT 8
US-09-895-814-879
/ Sequence 879, Application US/09895814
/ Publication No. US20020193296A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Dillon, Devin C.
/ APPLICANT: Mitchell, Jennifer L.
/ APPLICANT: Harlocker, Susan L.
/ APPLICANT: Ulang, Yugu
/ APPLICANT: Ralston, Michael D.
/ APPLICANT: Houghton, Robert A.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darriek
/ APPLICANT: Wang, Aijun X.
/ APPLICANT: Skeiky, Yashir A.W.
/ APPLICANT: Hepler, William T.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Hural, John
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Houghton, Raymond L.

```

```

/ APPLICANT: Vinals de Baseola, Carlota
/ APPLICANT: Foy, Teresa
/ APPLICANT: Fanger, Gary R.
/ TITLE OF INVENTION: LONG AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.42TC6
/ CURRENT APPLICATION NUMBER: US/09/895,814
/ CURRENT FILING DATE: 2001-06-29
/ NUMBER OF SEQ ID NOS: 990
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 879
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-09-895-814-879

Query Match      100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDINQFWTPTPTF 15
        |||||
Db      277 WIDINQFWTPTPTF 291

RESULT 9
US-10-010-667a-2
Sequence 2, Application US/10011095
Publication No. US20030045682A1
GENERAL INFORMATION:
/ APPLICANT: Afar, Daniel
/ APPLICANT: Leong, Rene S.
/ APPLICANT: Sateno, Arthur B.
/ APPLICANT: Saito, Shiro C.
/ APPLICANT: Mitchell, Steve Chappell
/ TITLE OF INVENTION: ANTIBODIES IMMUNOSPECIFIC FOR STAP1 (AS AMENDED)
/ FILE REFERENCE: 511582001610
/ CURRENT APPLICATION NUMBER: US/10/011,095
/ CURRENT FILING DATE: 2001-12-06
/ PRIOR APPLICATION NUMBER: 09/323,873
/ PRIOR FILING DATE: 1999-06-01
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/091,183
/ PRIOR FILING DATE: 1998-06-30
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ TYPE: PRF
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: DNA
US-10-011-095-2

Query Match      100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDINQFWTPTPTF 15
        |||||
Db      277 WIDINQFWTPTPTF 291

```

```

/ APPLICANT: Raitano, Arthur B.
/ APPLICANT: Saitan, Douglas C.
/ APPLICANT: Mitchell, Steve Chappell
/ TITLE OF INVENTION: NOVEL SEQUENCE TRANSMEMBRANE ANTIGENS
/ FILE REFERENCE: 511582001601
/ CURRENT APPLICATION NUMBER: US/10/010,667A
/ CURRENT FILING DATE: 2001-12-06
/ PRIOR APPLICATION NUMBER: 09/323,873
/ PRIOR FILING DATE: 1999-06-01
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/091,183
/ PRIOR FILING DATE: 1998-06-30
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-10-010-667a-2

Query Match      100.0%; Score 93; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDINQFWTPTPTF 15
        |||||
Db      277 WIDINQFWTPTPTF 291

RESULT 11
US-09-759-143-879
Sequence 2, Application US/09759143
Publication No. US2002002248A1
GENERAL INFORMATION:
/ APPLICANT: Xu, Jianshun
/ APPLICANT: Dillon, David C.
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: Harlocker, Susan L.
/ APPLICANT: Jiang, Inqui
/ APPLICANT: Harlocker, Susan L.
/ APPLICANT: Kallor, Michael D.
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Patrick
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skelley, Yaser A.W.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.42TC3
/ CURRENT APPLICATION NUMBER: US/09/759,143
/ CURRENT FILING DATE: 2001-01-12
/ NUMBER OF SEQ ID NOS: 934
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 879
/ LENGTH: 339
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-09-759-143-879

Query Match      100.0%; Score 93; DB 10; Length 339;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDINQFWTPTPTF 15
        |||||
Db      277 WIDINQFWTPTPTF 291

```

```

RESULT 12
US-09-780-669-879
/ Sequence 879; Application US/09780669
/ Patent No. US20020051977A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Dillon, Devin C.
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: Jia, Yueshan L.
/ APPLICANT: Jiang Yu, Susan L.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Seidel, Thomas S.
/ APPLICANT: Carter, Patrick
/ APPLICANT: Li, Samuel
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skelvy, Yaeir A.W.
/ APPLICANT: Hepler, William
/ APPLICANT: Hural, John
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Hershenson, John
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.427C24
/ CURRENT FILING DATE: 2001-02-09
/ NUMBER OF SEQ ID NOS: 943
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO: 879
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Homo sapiens

Query Match      100.0%; Score 93; DB 10; Length 339;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFWMTPTPF 15
Db      277 WIDIKQFWMTPTPF 291

RESULT 13
US-09-822-827-879
/ Sequence 879; Application US/09822827
/ Patent No. US20020081680A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.534C1US/09/822.827
/ CURRENT FILING DATE: 2001-03-28
/ NUMBER OF SEQ ID NOS: 982
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO: 879
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Homo sapiens

Query Match      100.0%; Score 93; DB 10; Length 339;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDIKQFWMTPTPF 15
Db      277 WIDIKQFWMTPTPF 291

```

```

Db      277 WIDIKQFWMTPTPF 291

RESULT 14
US-09-854-816-39
/ Sequence 39; Application US/09854816
/ Patent No. US20020151473A1
/ GENERAL INFORMATION:
/ APPLICANT: Andrew C. Braisted
/ APPLICANT: J. Christopher Phelan
/ APPLICANT: Robert S. McDowell
/ APPLICANT: Melissa A. Starovannik
/ APPLICANT: James A. Wells
/ TITLE OF INVENTION: Constrained Helical Peptides and Methods of
/ Making Same
/ NUMBER OF SEQUENCES: 113
/ CORRESPONDING PUBLICATION:
/ ADDRESSER: Genentech, Inc.
/ STREET: 1 DNA Way
/ CITY: South San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94080
/ COMPUTER READABLE FORM:
/ FILE TYPE: 1.44 Mb floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WinPatIn (Genentech)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/854,816
/ FILING DATE: 15-May-2001
/ CLASSIFICATION: <Unknown>
/ PRIORITY:
/ PRIOR APPLICATION NUMBER: 08/965,056
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Torchia, PhD., Timothy E.
/ REGISTRATION NUMBER: 36,700
/ REFERENCE/DOCKET NUMBER: P100582
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650/952-8879
/ TELEFAX: 650/952-8881
/ INFORMATION FOR SEQ ID NO: 39:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 268 amino acids
/ TYPE: Amino Acid
/ TOPOLOGY: Linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-854-816-39

Query Match      48.4%; Score 45; DB 10; Length 268;
Best Local Similarity 60.0%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 WIDIKQFWY 10
Db      202 WFDIOWMT 211

RESULT 15
US-09-854-816-33
/ Sequence 33; Application US/09854816
/ Patent No. US20020151473A1
/ GENERAL INFORMATION:
/ APPLICANT: Andrew C. Braisted
/ APPLICANT: J. Kevin Judice
/ APPLICANT: Robert S. McDowell
/ APPLICANT: J. Christopher Phelan
/ APPLICANT: Melissa A. Starovannik
/ APPLICANT: James A. Wells
/ TITLE OF INVENTION: Constrained Helical Peptides and Methods of

```

```

/      NUMBER OF SEQUENCES: 113      Making Same
/      CORRESPONDENCE ADDRESS:
/      ADDRESS: Genentech, Inc.
/      STREET: DNA
/      CITY: South San Francisco
/      STATE: California
/      COUNTRY: USA
/      ZIP: 94080
/      COMPUTER READABLE FORM:
/      MEDIA TYPE: 3.5 inch, 1.44 Mb floppy disk
/      OPERATING SYSTEM: PC-DOS/MS-DOS
/      SOFTWARE: Minipatn (Genentech)
/      CURRENT APPLICATION DATA:
/      APPLICATION NUMBER: US/09/854,816
/      FILING DATE: 15-May-2001
/      CLASSIFICATION: <Unknown>
/      PRIOR APPLICATION DATA:
/      APPLICATION NUMBER: 08/965,056
/      FILING DATE: <Unknown>
/      ATTORNEY/AGENT INFORMATION:
/      NAME: Torchia, PH.D., Timothy E.
/      REGISTRATION NUMBER: 36,700
/      TELECOMMUNICATION INFORMATION:
/      TELEPHONE: 650/222-8774
/      TELEFAX: 650/222-8881
/      INFORMATION FOR SEQ ID NO: 33:
/      SEQUENCE CHARACTERISTICS:
/      LENGTH: 269 amino acids
/      TYPE: Amino Acid
/      TOPOLOGY: Linear
/      SEQUENCE DESCRIPTION: SEQ ID NO: 33:
/      US-09-854-816-33
/
Query Match      48.4%; Score 45; DB 10; Length 269;
Best Local Similarity 60.0%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 0; Gaps 0;
QY      1 WIDKQPVWY 10
      1111111111
      203 WPDITWMT 212
DB

```

Search completed: March 26, 2003, 17:04:17
 Job time : 11.1163 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Comphen Ltd.

OM protein - protein search, using SW model

Run on: March 26, 2003, 16:44:40 / Search time 9.59302 Seconds
(Without alignments)
46.007 Million cell updates/sec

Title: US-10-010-667a-21

Sequence: 93

Sequence: 1 WIDIQQFWYTPPTF 15

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Issued Patents AA:*

1: /cgn2_6/prodata/1/aa/5A.COMB.pep:*

2: /cgn2_6/prodata/1/aa/5B.COMB.pep:*

3: /cgn2_6/prodata/1/aa/5A.COMB.pep:*

4: /cgn2_6/prodata/1/aa/5B.COMB.pep:*

5: /cgn2_6/prodata/1/aa/5A.COMB.pep:*

6: /cgn2_6/prodata/1/aa/5B.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	93	100.0	15	US-09-323-873A-21	Sequence 21, App1
2	93	100.0	15	US-09-323-873A-21	Sequence 2, App1
3	45	48.4	268	US-08-965-056-39	Sequence 39, App1
4	45	48.4	268	US-08-965-056-39	Sequence 3, App1
5	45	48.4	269	US-08-965-056-39	Sequence 45, App1
6	45	48.4	353	US-08-965-056-39	Sequence 45, App1
7	45	48.4	353	US-08-965-056-39	Sequence 11, App1
8	45	48.4	865	US-07-965-483-13	Sequence 4, App1
9	44	47.3	403	US-08-887-534A-83	Sequence 83, App1
10	43	46.2	351	US-08-817-441-47	Sequence 47, App1
11	43	46.2	873	US-08-912-129A-61	Sequence 61, App1
12	43	46.2	877	US-08-817-441-102	Sequence 102, App1
13	43	46.2	268	US-08-965-056-26	Sequence 48, App1
14	42	45.2	268	US-08-965-056-71	Sequence 71, App1
15	42	45.2	268	US-08-965-056-71	Sequence 19, App1
16	41.5	44.6	231	US-08-448-489-19	Sequence 642, App1
17	41	44.1	237	US-08-388-353-642	Sequence 642, App1
18	41	44.1	237	US-08-488-551B-642	Sequence 6, App1
19	41	44.1	243	US-08-965-056-73	Sequence 8, App1
20	41	44.1	268	US-08-965-056-8	Sequence 8, App1
21	41	44.1	268	US-08-965-056-11	Sequence 10, App1
22	41	44.1	268	US-08-965-056-10	Sequence 70, App1
23	41	44.1	268	US-08-965-056-10	Sequence 72, App1
24	41	44.1	269	US-08-965-056-12	Sequence 21, App1
25	41	44.1	269	US-08-965-056-11	Sequence 22, App1
26	41	44.1	269	US-08-965-056-25	Sequence 25, App1

28	41	44.1	269	US-08-965-056-37	Sequence 37, App1
29	41	44.1	270	US-08-965-056-67	Sequence 67, App1
30	41	44.1	351	US-08-470-770-46	Sequence 46, App1
31	41	44.1	351	US-08-468-059-46	Sequence 46, App1
32	41	44.1	351	US-09-109-916-46	Sequence 46, App1
33	41	44.1	351	US-09-109-916-46	Sequence 15, App1
34	41	44.1	602	US-09-257-490-15	Sequence 28, App1
35	41	44.1	850	US-08-448-603A-28	Sequence 28, App1
36	41	44.1	850	US-09-134-075-28	Sequence 28, App1
37	41	44.1	850	US-09-492-739-28	Sequence 28, App1
38	41	44.1	850	US-08-348-803-6	Sequence 6, App1
39	41	44.1	855	US-08-647-714-6	Sequence 6, App1
40	41	44.1	855	US-08-976-285-14	Sequence 14, App1
41	41	44.1	1503	US-08-976-285-10	Sequence 10, App1
42	40.5	43.5	848	US-09-083-521-7	Sequence 7, App1
43	40.5	43.5	1317	US-09-433-428D-67	Sequence 11, App1
44	40.5	43.5	1384	US-09-433-428D-67	Sequence 67, App1
45	40	43.0	149	US-09-433-428D-67	Sequence 67, App1

ALIGNMENTS

```

RESULT 1
US-09-323-873A-21
/ Sequence 21, Application US/09323873A
/ US-09-323-873A-21
/ GENERAL INFORMATION:
/ APPLICANT: Daniel E. Afar
/ APPLICANT: Rene S. Hubert
/ APPLICANT: Arthur B. Raitano
/ APPLICANT: Douglas C. Saffran
/ APPLICANT: Steve Chappell Mitchell
/ TITLE OF INVENTION: HUMAN TRANSMEMBRANE ANTIGENS
/ FILE REFERENCE: 129.16US02
/ CURRENT APPLICATION NUMBER: US/09/323,873A
/ PRIOR FILING DATE: 1999-06-01
/ PRIOR APPLICATION NUMBER: 60/091,183
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 21
/ LENGTH: 15
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURES:
/ INFORMATION: STBAP-1-PEPTIDE
US-09-323-873A-21
Query Match 100.0%; Score 93; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 2,4e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 WIDIQQFWYTPPTF 15
US-09-323-873A-21
/ Sequence 2, Application US/09323873A
/ US-09-323-873A-21
/ Patent No. 6,249,500
/ ORGANISM: HUMAN
/ APPLICANT: Daniel E. Afar
/ APPLICANT: Rene S. Hubert
/ APPLICANT: Arthur B. Raitano
/ APPLICANT: Douglas C. Saffran
/ APPLICANT: Steve Chappell Mitchell

```

```

/ TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
/ TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USSES THEREOF
/ CURRENT REFERENCE: 129,160S12
/ CURRENT APPLICATION NUMBER: US/09/323,873A
/ PRIORITY DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/087,520
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/091,183
/ PRIOR FILING DATE: 1998-06-30
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO: 2
/ SEQ ID NO: 3
/ SEQ ID NO: 4
/ TYPE: PR1
/ ORGANISM: Homo Sapiens
/ US-09-323-873A-2

Query Match      100.0%; Score 93; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 5,6e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WIDKQFWMTPTPF 15
Db      277 WIDKQFWMTPTPF 291

RESULT 3
US-08-965-056-39
/ Sequence 39, Application US/08965056
/ Patent No. 6271198
/ GENERAL INFORMATION:
/ APPLICANT: Andrew C. Braisted
/ APPLICANT: J. Kevin Judice
/ APPLICANT: Robert S. McDowell
/ APPLICANT: J. Christopher Phelan
/ APPLICANT: James A. Starovastnik
/ APPLICANT: James A. Wells
/ TITLE OF INVENTION: Constrained Helical Peptides and Methods of
/ TITLE OF INVENTION: Making Same
/ NUMBER OF SEQUENCES: 113
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Genentech, Inc.
/ STREET: 460 Point San Bruno Blvd
/ CITY: South San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94080
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WinPatIn (Genentech)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/965,056
/ FILING DATE: 05-No. 6271198-1997
/ CLASSIFICATION: 530
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Torchia, PhD., Timothy E.
/ REGISTRATION NUMBER: 36,700
/ REFERENCE/DOCKET NUMBER: P100582
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650/225-8674
/ TELEFAX: 650/952-9881
/ INFORMATION FOR SEQ ID NO: 39:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 269 amino acids
/ TYPE: Amino Acid
/ TOPOLOGY: Linear
/ US-08-965-056-39

Query Match      48.4%; Score 45; DB 4; Length 269;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 WIDKQFWMT 10
Db      202 WFDLQWMT 211

RESULT 4
US-08-965-056-33
/ Sequence 33, Application US/08965056
/ Patent No. 6271198
/ GENERAL INFORMATION:
/ APPLICANT: Andrew C. Braisted
/ APPLICANT: J. Kevin Judice
/ APPLICANT: Robert S. McDowell
/ APPLICANT: J. Christopher Phelan
/ APPLICANT: Melissa A. Starovastnik
/ APPLICANT: James A. Wells
/ TITLE OF INVENTION: Constrained Helical Peptides and Methods of
/ TITLE OF INVENTION: Making Same
/ NUMBER OF SEQUENCES: 113
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Genentech, Inc.
/ STREET: 1 DNA Way
/ CITY: South San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94080
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WinPatIn (Genentech)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/965,056
/ FILING DATE: 05-No. 6271198-1997
/ CLASSIFICATION: 530
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Torchia, PhD., Timothy E.
/ REGISTRATION NUMBER: 36,700
/ REFERENCE/DOCKET NUMBER: P100582
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650/225-8674
/ TELEFAX: 650/952-9881
/ INFORMATION FOR SEQ ID NO: 33:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 269 amino acids
/ TYPE: Amino Acid
/ TOPOLOGY: Linear
/ US-08-965-056-33

Query Match      48.4%; Score 45; DB 4; Length 269;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 WIDKQFWMT 10
Db      203 WFDLQWMT 212

RESULT 5
US-08-118-270-45
/ Sequence 45, Application US/08118270
/ Patent No. 5508384
/ GENERAL INFORMATION:
/ APPLICANT: Murphy, Randall B.
/ APPLICANT: Scott, Robert
/ APPLICANT: Schmitt, Robert
/ TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
/ TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
/ NUMBER OF SEQUENCES: 348
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: BROWDY AND NEITRAK
/ STREET: 419 Seventh Street, N.W., Suite 300
/ CITY: Washington

```

```

/ STATE: D.C.
/ COUNTRY: USA
/ SERIAL: 20004
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US-08/118,270
/ PRIORITY DATE: 09-SEP-1993
/ APPLICATION NUMBER: US 07/943,236
/ FILING DATE: 10-SEP-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Townsend, Kevin G.
/ REGISTRATION NUMBER: 34,033
/ REFERENCE/DOCKET NUMBER: MURPHY-2A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-737-3528
/ TELEFAX: 202-737-3528
/
/ INFORMATION FOR SEQ ID NO: 45:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 353 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/
/ US-08-118-270-45
/
/ Query Match 48.4%, Score 45, DB 1, Length 353;
/ Best Local Similarity 33.3%, Pred. No. 15;
/ Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
/
/ Oy 1 MIDIKQPVWYTPPTP 15
/ Db 56 WVELYNFIMHPMAF 70
/
/ RESULT 6
/ PCT-US93-08528-45
/ GENERAL INFORMATION:
/ APPLICANT: New York University
/ TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
/ NUMBER OF SEQUENCES: 348
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: BROWNY AND NEIMARK
/ STREET: 419 Seventh Street, N.W., Suite 300
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20004
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US93/08528
/ FILING DATE: 09-SEP-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/943,236
/ FILING DATE: 10-SEP-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Townsend, Kevin G.
/ REGISTRATION NUMBER: 34,033
/ REFERENCE/DOCKET NUMBER: MURPHY-2A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-628-5197
/ TELEFAX: 202-737-3528

```

```

/ TELELEX: 248633
/ INFORMATION FOR SEQ ID NO: 45:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 353 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/
/ PCT-US93-08528-45
/
/ Query Match 48.4%, Score 45, DB 5, Length 353;
/ Best Local Similarity 33.3%, Pred. No. 15;
/ Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
/
/ Oy 1 MIDIKQPVWYTPPTP 15
/ Db 56 WVELYNFIMHPMAF 70
/
/ RESULT 7
/ US-07-956-483-13
/ Sequence 13, Application US/07956483
/ Patent No. 6261799
/ GENERAL INFORMATION:
/ APPLICANT: KIENY, Marie-Paule
/ TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEANED
/ NUMBER OF SEQUENCES: 9240 VARIANT
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Burne, Deane, Swecker & Mathie
/ STREET: P.O. Box 1404
/ CITY: Alexandria
/ STATE: Virginia
/ COUNTRY: United States
/ ZIP: 22313-1404
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US-07/956,483
/ FILING DATE: 12-DEC-1992
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: NO 92/19742
/ FILING DATE: 12-NOV-1992
/ PRIOR APPLICATION DATA: FR 91 05392
/ APPLICATION NUMBER: FR 91 05392
/ FILING DATE: 02-MAY-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Crane-Feunty, Sharon E.
/ REGISTRATION NUMBER: 36,113
/ REFERENCE/DOCKET NUMBER: 017753-005
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-6620
/ TELEFAX: (703) 836-2021
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 865 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/
/ US-07-956-483-13
/
/ Query Match 48.4%, Score 45, DB 4, Length 865;
/ Best Local Similarity 60.0%, Pred. No. 36;
/ Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
/
/ Oy 1 MIDIKQPVWY 10
/ Db 681 WFDITQMLWY 690

```

```

RESULT 8
US-08-472-240A-4
Patent No. 6284248
GENERAL INFORMATION:
APPLICANT: KIENY, Marie-Paule
TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSER: HUMS DOHNE, SWECKER & MATHIS
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,240A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIORITY NUMBER: 0481
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35, 030
REFERENCE/DOCKET NUMBER: 017753-055
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
IMPROVATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 887 amino acids
TYPE: amino acid
STRANDNESS: not relevant
APPLICANT: HUMS DOHNE, SWECKER & MATHIS
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..854
US-08-472-240A-4

Query Match 48.44; Score 15; DB 4; Length 887;
Local Similarity 40.00; Percent 2; Mismatches
Matches 6; Conservative 2; Indels 0; Gaps 0;

QY 1 MDIKQPVWY 10
DB 703 WDIOTWLMW 712

RESULT 9
US-08-487-534A-83
Sequence 83, Application US/0887534A
Patent No. 6455323
GENERAL INFORMATION:
APPLICANT: Holden, David W.
TITLE OF INVENTION: ANTI-BACTERIAL METHODS AND MATERIALS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSER: Marshall, O'Toole, Gerheine, Murray & Bonn
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/887,534A
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: KLIN-LAMBERG, L., 193-67
REGISTRATION NUMBER: 33, 547
REFERENCE/DOCKET NUMBER: 28341/33996
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 442 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-887-534A-83

Query Match 47.34; Score 44; DB 4; Length 403;
Local Similarity 30.31; Percent 24;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 DIKQPVWTPPT 14
DB 305 DFKWTWTPPT 316

RESULT 10
US-08-817-441-47
Sequence 47, Application US/08817441
Patent No. 6392924
GENERAL INFORMATION:
APPLICANT: CHARREAU, PIERRE
APPLICANT: CLAVEL, FRANCOIS
APPLICANT: BOWEN, ANDREW
APPLICANT: GUYARD, DENISE
APPLICANT: MONTAGNIER, LUC
APPLICANT: DONCON DE SAINT-MARTIN, JACQUELINE
APPLICANT: COHEN, JAOUES
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR
VIRUS) AND ANTIBODIES
CORRESPONDENCE ADDRESS:
ADDRESSER: Plimagan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,441
FILING DATE: 20-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 95/01391
FILING DATE: 20-OCT-1994
PRIOR APPLICATION DATA:

```



```

/ APPLICATION NUMBER: FR 9502526
/ FILING DATE: 03-MAR-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: WEYERS, KEMELCH J.
/ REGISTRATION NUMBER: 25,146
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-408-4000
/ TELEFAX: 202-408-4400
/ INFORMATION FOR SEQ ID NO: 47:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 931 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein

Query Match 46.2%; Score 43; DB 4; Length 351;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MIDIKOPVWY 10
Db 160 WLDITKWLWY 169

RESULT 11
US-08-912-129A-61
/ Sequence 61, Application US/08912129A
/ Patent No. 5922533
/ GENERAL INFORMATION:
/ APPLICANT: VALLEARY, ANADRUZEIA S.
/ APPLICANT: HICKLEY, JOHN DE.
/ APPLICANT: HICKLEY, ROBERT K.
/ APPLICANT: HICKLEY, ROBERT K., JR.
/ APPLICANT: NECKLAUS, ELIZABETH A.
/ APPLICANT: GOLDEN, ALAN M.
/ APPLICANT: BRENNAN, CATHERINE A.
/ APPLICANT: DEVADE, SUSHIL G.
/ TITLE OF INVENTION: RAPID ASSAY FOR SIMULTANEOUS DETECTION AND DIFFERENTIATIO
/ NUMBER OF SEQUENCES: 89
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSES: 100 Abbott Park Road
/ CITY: Abbott Park
/ STATE: IL
/ COUNTRY: USA
/ ZIP: 60064-3500
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
/ OPERATING SYSTEM: MS-DOS (Windows 95)
/ SOFTWARE: Microsoft Word (ASCII format output)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/912,129A
/ FILING DATE: 15-AUG-1997
/ CLASSIFICATION: 436
/ PRIOR APPLICATION DATA:
/ PCT/FR 95/01391
/ FILING DATE: 20-OCT-1995
/ INFORMATION FOR SEQ ID NO: 61:
/ NAME: Danckers, Andreas M.
/ REGISTRATION NUMBER: 32,652
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-408-4000
/ TELEFAX: 202-408-4400
/ INFORMATION FOR SEQ ID NO: 61:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 873 amino acids
/ TYPE: amino acid

```

```

/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: Protein
/ US-08-912-129A-61

Query Match 46.2%; Score 43; DB 2; Length 873;
Best Local Similarity 50.0%; Pred. No. 75;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MIDIKOPVWY 10
Db 682 WLDITKWLWY 691

RESULT 12
US-08-817-441-102
/ Sequence 102, Application US/08817441
/ Patent No. 639234
/ GENERAL INFORMATION:
/ APPLICANT: JACQUES, PIERRE
/ APPLICANT: CLAVEL, FRANCOIS
/ APPLICANT: BORMAN, ANDREW
/ APPLICANT: OUILLET, CAROLINE
/ APPLICANT: GUERARD, DENISE
/ APPLICANT: MONTAGNIER, LUC
/ APPLICANT: DONJON DE SAINT-MARTIN, JACQUELINE
/ APPLICANT: JOHNS, JAMES ELOTIDE SPOURCES OF HIV-1 TYPE (OR
/ TITLE OF INVENTION: SURVEYED ANTIGENS
/ NUMBER OF SEQUENCES: 103
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSES: Flanagan, Faradow, Garrett &
/ ADDRESSER: Dunner, L.L.P.
/ STREET: 1300 I Street, N.W.
/ CITY: Washington
/ STATE: Washington
/ COUNTRY: USA
/ ZIP: 20005-3315
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Microsoft Word 6.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/817,441
/ FILING DATE: 11-JUL-1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/FR 95/01391
/ FILING DATE: 20-OCT-1995
/ INFORMATION FOR SEQ ID NO: 102:
/ NAME: Weyers, Kemelch J.
/ REGISTRATION NUMBER: 25,146
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-408-4000
/ TELEFAX: 202-408-4400
/ INFORMATION FOR SEQ ID NO: 102:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 877 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein

Query Match 46.2%; Score 43; DB 4; Length 877;
Best Local Similarity 50.0%; Pred. No. 75;

```

Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WIDIKQPMWT 10
 Db 666 WIDISRWMLWT 695

RESULT 13
 US-08-965-056-26
 Sequence 26, Application US/08965056
 Patent No. 6271198
 GENERAL INFORMATION:
 APPLICANT: Andrew C. Braisted
 APPLICANT: J. Kevin Judice
 APPLICANT: Robert S. McDowell
 APPLICANT: J. Christopher Phelan
 APPLICANT: Melissa A. Starovasilk
 TITLE OF INVENTION: Constrained helical peptides and Methods of
 MAKING SAME
 NUMBER OF SEQUENCES: 113
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 1 DNA Way
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WinPacIn (Genentech)
 CURRENT APPLICATION DATA:
 FILING DATE: 05-NO. 6271198-1997
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Torchia, PhD., Timothy E.
 REGISTRATION NUMBER: 36,700
 REFERENCE/DOCKET NUMBER: P100582
 TELEPHONE: 650/225-8674
 TELEFAX: 650/952-9881
 INFORMATION FOR SEQ ID NO: 26:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 268 amino acids
 TYPE: Amino Acid
 TOPOLOGY: Linear
 US-08-965-056-26

Query Match 45.2%; Score 42; DB 4; Length 268;
 Best Local Similarity 50.0%; Pred. No. 32;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WIDIKQPMWT 10
 Db 202 WIDISRWMLWT 211

RESULT 14
 US-08-965-056-68
 Sequence 68, Application US/08965056
 Patent No. 6271198
 GENERAL INFORMATION:
 APPLICANT: Andrew C. Braisted
 APPLICANT: J. Kevin Judice
 APPLICANT: Robert S. McDowell
 APPLICANT: J. Christopher Phelan
 APPLICANT: Melissa A. Starovasilk
 TITLE OF INVENTION: Constrained helical peptides and Methods of
 MAKING SAME

NUMBER OF SEQUENCES: 113
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 1 DNA Way
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WinPacIn (Genentech)
 CURRENT APPLICATION DATA:
 FILING DATE: 05-NO. 6271198-1997
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Torchia, PhD., Timothy E.
 REGISTRATION NUMBER: 36,700
 REFERENCE/DOCKET NUMBER: P100582
 TELEPHONE: 650/225-8674
 TELEFAX: 650/952-9881
 INFORMATION FOR SEQ ID NO: 68:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 268 amino acids
 TYPE: Amino Acid
 TOPOLOGY: Linear
 US-08-965-056-68

Query Match 45.2%; Score 42; DB 4; Length 268;
 Best Local Similarity 50.0%; Pred. No. 32;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WIDIKQPMWT 10
 Db 202 WIDISRWMLWT 211

RESULT 15
 US-08-965-056-71
 Sequence 71, Application US/08965056
 Patent No. 6271198
 GENERAL INFORMATION:
 APPLICANT: Andrew C. Braisted
 APPLICANT: J. Kevin Judice
 APPLICANT: Robert S. McDowell
 APPLICANT: J. Christopher Phelan
 APPLICANT: Melissa A. Starovasilk
 TITLE OF INVENTION: Constrained helical peptides and Methods of
 MAKING SAME
 NUMBER OF SEQUENCES: 113
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 1 DNA Way
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WinPacIn (Genentech)
 CURRENT APPLICATION DATA:
 FILING DATE: 05-NO. 6271198-1997
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Torchia, PhD., Timothy E.
 REGISTRATION NUMBER: 36,700

```

/ REFERENCE/DOCKET NUMBER: P1005R2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650/4228674
/ FAX: 650/4228674
/ INFORMATION FOR SEQ ID NO: 71:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 268 amino acids
/ TYPE: Amino Acid
/ TOPOLOGY: Linear
/ US-08-965-056-71

Query Match 45 24; Score 42; DB 4; Length 268;
Best Local Similarity 50.04; Pred. No. 32;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDIKQFVWY 10
DB 202 WPDISRRLWY 211

```

Search completed: March 26, 2003, 16:51:41
Job time : 10.593 secs

Genome version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: March 26, 2003, 16:43:45 ; Search time 15.6977 Seconds
(without alignments)
196,889 Million cell updates/sec

Title: US-10-010-667a-21
Perfect score: 93

Sequence: 1 WIDKQFWYTPPTF 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapeop 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_plant:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_ricent:*
12: sp_virus:*
13: sp_viridate:*
14: sp_viridate:*
15: sp_viridate:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	97.8	338	1	O9GL50 sus scrota
2	80	86.0	338	1	O9GL50 sus musculus
3	80	86.0	338	1	O9GL50 sus musculus
4	80	86.0	338	1	O9GL50 sus musculus
5	49	52.7	448	16	O9GL50 sus musculus
6	48	51.6	453	16	O9GL50 sus musculus
7	46	49.5	122	15	O9GL50 sus musculus
8	46	49.5	122	15	O9GL50 sus musculus
9	46	49.5	122	15	O9GL50 sus musculus
10	46	49.5	122	15	O9GL50 sus musculus
11	46	49.5	122	15	O9GL50 sus musculus
12	46	49.5	122	15	O9GL50 sus musculus
13	46	49.5	122	15	O9GL50 sus musculus
14	46	49.5	122	15	O9GL50 sus musculus
15	46	49.5	122	15	O9GL50 sus musculus
16	46	49.5	122	15	O9GL50 sus musculus

17	45	48.4	133	15	O9GL50 sus musculus
18	45	48.4	133	15	O9GL50 sus musculus
19	45	48.4	133	15	O9GL50 sus musculus
20	45	48.4	133	15	O9GL50 sus musculus
21	45	48.4	133	15	O9GL50 sus musculus
22	45	48.4	133	15	O9GL50 sus musculus
23	45	48.4	133	15	O9GL50 sus musculus
24	45	48.4	133	15	O9GL50 sus musculus
25	45	48.4	133	15	O9GL50 sus musculus
26	45	48.4	133	15	O9GL50 sus musculus
27	45	48.4	133	15	O9GL50 sus musculus
28	45	48.4	133	15	O9GL50 sus musculus
29	45	48.4	133	15	O9GL50 sus musculus
30	45	48.4	133	15	O9GL50 sus musculus
31	45	48.4	133	15	O9GL50 sus musculus
32	45	48.4	133	15	O9GL50 sus musculus
33	45	48.4	133	15	O9GL50 sus musculus
34	45	48.4	133	15	O9GL50 sus musculus
35	45	48.4	133	15	O9GL50 sus musculus
36	45	48.4	133	15	O9GL50 sus musculus
37	45	48.4	133	15	O9GL50 sus musculus
38	45	48.4	133	15	O9GL50 sus musculus
39	45	48.4	133	15	O9GL50 sus musculus
40	45	48.4	133	15	O9GL50 sus musculus
41	45	48.4	133	15	O9GL50 sus musculus
42	45	48.4	133	15	O9GL50 sus musculus
43	45	48.4	133	15	O9GL50 sus musculus
44	45	48.4	133	15	O9GL50 sus musculus
45	45	48.4	133	15	O9GL50 sus musculus

ALIGNMENTS

RESULT 1	O9GL50	PRELIMINARY;	PR1:	338 AA.
ID	O9GL50			
AC	O9GL50			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)			
DE	Six transmembrane endothelial antigen of PACC.			
OS	Human (Homo sapiens)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; NCBI_TaxId=9623;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Nagasaka T., Bouillon G., Coupel S., Coulon F., Tesson L., Heston J.-W., Souillon J.-P., Charreau B.;			
RT	Transmembrane gene expression in endothelial cells during TNF-alpha-			
RT	induced endothelial cell apoptosis			
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF139659; ANJ3868.1; -;			
KW	Transmembrane.			
SQ	SEQUENCE 338 AA; 39918 MW; ED4908E067A32B CRC64;			
Query Match	97.8%; Score 91; DB 6; Length 338;			
Best Local Similarity	86.7%; Pred. No. 3.4e-07;			
Matches	13; Conservative 2; Mismatches 0; Indels 0; Gaps 0;			
DB	1 WIDKQFWYTPPTF 15			
276	WIDKQFWYTPPTF 290			
RESULT 2	O9GL50	PRELIMINARY;	PR1:	339 AA.
ID	O9GL50			
AC	O9GL50			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			

```

DE 241007B1SR1K protein.
GN STEAP OR 241007B1SR1K.
OS Mus musculus (mouse); Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRIONIC STEM CELLS;
RA MEDLINE=108560; PubMed=1121781;
RA Kawai T., Shingawa K., Shikama K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasuza T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fiedelmann M., Gaesteland T., Gissi C., King B., Kochiya H.,
RA Knehl P., Lewis S., Matsuo Y., Mikado I., Pesole G., Quackenbush J.,
RA Schmitt L., Scudlitz F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Boffelli D., Bouhassira D., Brown J., de Bonaldo M.P.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.P.,
RA Garschütz S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Norioka P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki M., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki K., Toyooka K., Wang K.H., Wetz C., Wittaker C., Wilmink U.,
RA Hayashizaki Y., Itohida K., Hasegawa I., Kameji H., Kohzuki S.,
RA "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK010437; BAE26938.1; -.
SQ SEQUENCE 339 AA; 39264 MW; 3F7AB9C7520F0966 CRC64;

Query Match 86.0%; Score 80; DB 11; Length 339;
Best Local Similarity 73.3%; Pred. No. 2,4e-05;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDKQFWMTPTPT 15
DB 277 WVDVSGFWMTPTPT 291
|||||
RESULT 3
Q92422 PRELIMINARY; PRT; 339 AA.
AC Q92422.1 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Dnublin-2002 (TEMBLrel. 21, Last annotation update)
GN 101001D01SR1K.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRIONIC STEM CELLS;
RA MEDLINE=108560; PubMed=1121781;
RA Kawai T., Shingawa K., Shikama K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasuza T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fiedelmann M., Gaesteland T., Gissi C., King B., Kochiya H.,
RA Knehl P., Lewis S., Matsuo Y., Mikado I., Pesole G., Quackenbush J.,
RA Schmitt L., Scudlitz F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Boffelli D., Bouhassira D., Brown J., de Bonaldo M.P.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.P.,
RA Garschütz S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Norioka P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki M., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki K., Toyooka K., Wang K.H., Wetz C., Wittaker C., Wilmink U.,
RA Hayashizaki Y., Itohida K., Hasegawa I., Kameji H., Kohzuki S.,
RA "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK010437; BAE26938.1; -.
SQ SEQUENCE 339 AA; 39264 MW; 3F7AB9C7520F0966 CRC64;

Query Match 86.0%; Score 80; DB 11; Length 339;
Best Local Similarity 73.3%; Pred. No. 2,4e-05;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDKQFWMTPTPT 15
DB 277 WVDVSGFWMTPTPT 291
|||||

```

```

DB 277 WVDVSGFWMTPTPT 291
|||||
RESULT 4
Q92429 PRELIMINARY; PRT; 339 AA.
AC Q92429.1 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
GN STEAP OR 241007B1SR1K.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRIONIC STEM CELLS;
RA MEDLINE=108560; PubMed=1121781;
RA Kawai T., Shingawa K., Shikama K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasuza T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fiedelmann M., Gaesteland T., Gissi C., King B., Kochiya H.,
RA Knehl P., Lewis S., Matsuo Y., Mikado I., Pesole G., Quackenbush J.,
RA Schmitt L., Scudlitz F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Boffelli D., Bouhassira D., Brown J., de Bonaldo M.P.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.P.,
RA Garschütz S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Norioka P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki M., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki K., Toyooka K., Wang K.H., Wetz C., Wittaker C., Wilmink U.,
RA Hayashizaki Y., Itohida K., Hasegawa I., Kameji H., Kohzuki S.,
RA "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK010437; BAE26938.1; -.
SQ SEQUENCE 339 AA; 39264 MW; 3F7AB9C7520F0966 CRC64;

Query Match 86.0%; Score 80; DB 11; Length 339;
Best Local Similarity 73.3%; Pred. No. 2,4e-05;
Matches 11; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 WIDKQFWMTPTPT 15
DB 277 WVDVSGFWMTPTPT 291
|||||
RESULT 5
Q92429 PRELIMINARY; PRT; 448 AA.
AC Q92429.1 (TEMBLrel. 18, Created)
DT 01-OCT-2001 (TEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TEMBLrel. 18, Last annotation update)
DE 01-OCT-2001 (TEMBLrel. 18, Last annotation update)
DE Alpha-glucosidase ABC transporter, substrate-binding protein.
GN Hs2113.1. loci (Memorized loci)
OS Homo sapiens (human).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobiaceae; Neorhizobiaceae.
NCBI_TaxID=381.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NAF030309; PubMed=11214968;
RA MEDLINE=108560; PubMed=11214968;
RA Kawai T., Shingawa K., Shikama K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasuza T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fiedelmann M., Gaesteland T., Gissi C., King B., Kochiya H.,
RA Knehl P., Lewis S., Matsuo Y., Mikado I., Pesole G., Quackenbush J.,
RA Schmitt L., Scudlitz F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Boffelli D., Bouhassira D., Brown J., de Bonaldo M.P.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.P.,
RA Garschütz S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Norioka P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki M., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki K., Toyooka K., Wang K.H., Wetz C., Wittaker C., Wilmink U.,
RA Hayashizaki Y., Itohida K., Hasegawa I., Kameji H., Kohzuki S.,
RA "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK010437; BAE26938.1; -.
SQ SEQUENCE 448 AA; 48920 MW; 9A7B8569A6945D19 CRC64;

Query Match 52.7%; Score 49; DB 16; Length 448;
Best Local Similarity 53.8%; Pred. No. 4.3;

```

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 DIKQFWTPPF 15
DB 164 DVALSWVBNPF 176

RESULT 6
OSUHT8 PRELIMINARY; PRT; 453 AA.

ID OSUHT8
AC OSUHT8, 2002 (TRENBIrel, 21, Created)
DT 01-JUN-2002 (TRENBIrel, 21, Last sequence update)
DE ABC transporter, substrate binding protein.
GN AGC OR ATU0591 OR AGC_C_1045.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
NCBI_TaxID=16299;
DB 164 DVALSWVBNPF 176

SEQUENCE FROM N.A.
RX MEDLINE-21608550; PubMed-11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kiteajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Mod G.B., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D., Sr.,
RA Chapman P., Glendinning Y., Deatherage G., Gillet M., Grant C.,
RA Karp P.D., Karp M., Karp L., Karp A., Karp Z., Romero P.,
RA Raymond C., Rouse G., Samplinschak C.-W. Z., Romero P., Perry M.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.,
RT The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.
RN [2] since 294.12317-2223 (2001).

SEQUENCE FROM N.A.
RX MEDLINE-21608551; PubMed-11743194;
RA Goodner B., Hinkle G., Gatlung S., Miller N., Blanchard M.,
RA Quiclio B., Goldman B.S., Cao Y., Askenzai M., Halling C., Mullin L.,
RA Hummel K., Gordon J., Vaudin M., Iarchoux O., Epp A., Liu F.,
RA Platanus C., Anguel M., Doughty D., Scott C., Kappes B.,
RA Cielo C., Slater S. J., Gustin J., Lomo C., Sest C., Strub G.,
RT Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.
RL Science 294.2323-2328 (2001).
DR EMBL; A8009027; AAL41608.1;
NCBI_TaxID=16299; AAB6402.1;
SO SPOONCR 453 AA; 49090 MW; 90896249313085F CRC64;

Query Match 51.6%; Score 48; DB 16; Length 453;
Best Local Similarity 53.8%; Pred. No. 7.2;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 DIKQFWTPPF 15
DB 169 DVALSWVBNPF 181

RESULT 7
OSUJL9 PRELIMINARY; PRT; 122 AA.

ID OSUJL9
AC OSUJL9, 2000 (TRENBIrel, 15, Created)
DT 01-OCT-2000 (TRENBIrel, 15, Last sequence update)
DE Env. envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MIDIKQFWY 10
DB 660 WFDISQWLWY 669

OSUHT8 PRELIMINARY; PRT; 453 AA.

ID OSUHT8
AC OSUHT8, 2002 (TRENBIrel, 21, Created)
DT 01-JUN-2002 (TRENBIrel, 21, Last sequence update)
DE ABC transporter, substrate binding protein.
GN AGC OR ATU0591 OR AGC_C_1045.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
NCBI_TaxID=16299;
DB 164 DVALSWVBNPF 176

SEQUENCE FROM N.A.
RX MEDLINE-21608550; PubMed-11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kiteajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Mod G.B., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D., Sr.,
RA Chapman P., Glendinning Y., Deatherage G., Gillet M., Grant C.,
RA Karp P.D., Karp M., Karp L., Karp A., Karp Z., Romero P.,
RA Raymond C., Rouse G., Samplinschak C.-W. Z., Romero P., Perry M.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.,
RT The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.
RN [2] since 294.12317-2223 (2001).

SEQUENCE FROM N.A.
RX MEDLINE-21608551; PubMed-11743194;
RA Goodner B., Hinkle G., Gatlung S., Miller N., Blanchard M.,
RA Quiclio B., Goldman B.S., Cao Y., Askenzai M., Halling C., Mullin L.,
RA Hummel K., Gordon J., Vaudin M., Iarchoux O., Epp A., Liu F.,
RA Platanus C., Anguel M., Doughty D., Scott C., Kappes B.,
RA Cielo C., Slater S. J., Gustin J., Lomo C., Sest C., Strub G.,
RT Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.
RL Science 294.2323-2328 (2001).
DR EMBL; A8009027; AAL41608.1;
NCBI_TaxID=16299; AAB6402.1;
SO SPOONCR 453 AA; 49090 MW; 90896249313085F CRC64;

Query Match 51.6%; Score 48; DB 16; Length 453;
Best Local Similarity 53.8%; Pred. No. 7.2;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 DIKQFWTPPF 15
DB 169 DVALSWVBNPF 181

RESULT 7
OSUJL9 PRELIMINARY; PRT; 122 AA.

ID OSUJL9
AC OSUJL9, 2000 (TRENBIrel, 15, Created)
DT 01-OCT-2000 (TRENBIrel, 15, Last sequence update)
DE Env. envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

	Query Match	48.4%	Score 45	DB 2	Length 107
	Seq. Ident.	96.2%	Ident. M		
	Matches	7	Conservative	2	Mismatches
				Indels	Gaps
Qy	1 WIDIKOPVWTPPTP 15				
	- - :- - -				
Dv	22 WKNIKILAWGSDTF 36				

	RESULT	10
O9YXR4		
ID	PRELIMINARY;	PRT: 122 AA.
AC	GQYXR4	
CD	GQYXR4_1989 (TEMBHirel_10, Created)	
DT	01-NOV-1999 (EMBL; last annotation update)	
DR	01-DEC-2001 (TEMBHirel_15, Last annotation update)	
DE	Envelope glycoprotein immunodominant region (Fragment).	
GN	ENV.	
OS	Human immunodeficiency virus type 1.	
OC	Viruses; Retroviridae; Retrovirales; Lentivirinae.	
NCBI	TaxID=11676;	
LN	[1]	
R	SEQUENCE FROM N.A.	
RA	PERMANENT DEPOSIT	
RA	Tamari A., Swanson P.A., Devare S.G., Barro O.J., Savadara A., Costa L.V., Telles J.C., Brindley R., Schable C., Plenzieske D., Rayfield W.;	
RL	"HIV-1 subtypes among blood donors from Rio de Janeiro, Brazil."	
RT	EMBL; AF030039; AACG9291.1;"	
DR	InterPro: IPR000328; Env_GP41.	
FT	Fram. PR06317; GP41; 1.	
KX	Unpublished	
PT	NON_TER	122
ST	SEQUENCE	122 AA; 14739 MW; 40D47B9BC8B8C2 CRC64;
Query Match		48-44; Score 45; DB 15; Length 122;
Best Local Similarity		60.04; Pred. No. 5.7;
Matches	6; Conservative	2; Mismatches
		2; Indels
Oy	1 WDIFQPNW 10	
Db	112 WDITOWLAW 121	
RESULT	11	

	ID	G9YX01		PRELIMINARY;	FRT,	122 AA.
	AC	G9YX01,				
	PC	87-1999	(T-FEMBL),	10, [Created]		
	DT	01-MAY-1999	(TREMBL).	10, [Last sequence update]		
	DT	01-DEC-2001	(TREMBL),	19, [Last annotation update]		
	DE	Envelope glycoprotein immunodominant region (Fragment),				
	GN	ENV.				
	OS	Human Immunodeficiency virus type 1,				
	VS	virusess; Retroid viruses; Retroviridae; Lentivirus.				
	OX	NCOI_TaxID=11676;				
	DN	SEQUENCE FROM N.A.				
	RJ	SFRAIN-RJ96BHPQ42;				
	RA	Tanuri A., Swenson P.A., Devare S.G., Risco O.J., Saviedra A.,				
	RA	Costa L.U., Telles J.G., Brindetto R., Scabale C., Plentzack D.,				
	RA	Rayfield M.;				
	RL	"HY-I" subtypes among blood donors from Rio de Janeiro, Brazil,"				
	RL	Submitted (NOV-1997) to the EMBL/genbank/DBD databases.				
	DR	EMBL; AF010562; U079504.1 -				
	DR	NCBI; GenBank; F010562.1; U079504.1.				
	DR	RefSeq: PP00517; GP41, 1.				
	KM	Transmembrane.				
	FT	NON_TER	1			
	FT	NON_TER	122			
	SO	SEQUENCE	122 AA; 14792 NM; 705BB601A6BBD17 CRC64;			
	Query Match		46.4%; Score 45; DB 15; Length 122;			
	% Local Similarity		60.4%; RefSeq NM 5,7;			
	Matches	6; Conservative	2; Mismatches	2; Indels	0; Gaps	0;
Qy	1 MIDIKPWWY	10				
Db	112 WFDITWLMWY	121				

[illegible]


```

RESULT 13
ID Q90018 PRELIMINARY; PRT: 122 AA.
AC Q90018;
DT 01-MAY-2000 (TREMBLrel, 13, Created)
DT 01-MAY-2000 (TREMBLrel, 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel, 19, Last annotation update)
DR Envelope glycoprotein (Fragment).
GN ENV. Human immunodeficiency virus type 1
OC Viruses; Retrod viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GP41ERR23;
RA Caride E., Hertogs K., Larder B., Dehertogh P., Brindeiro R.,
RA Machado E., de Sa C.A.M., Eyre W., Passioni L.F.C., Meneses J.A.,
RA D'Aquila R.T. and others.
RT "Genotyping and phylogeny analysis of B and non-B human
RT immunodeficiency virus type 1 subtypes from patients under HAART."
RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF165556; AAF08501.1; -
DR InterPro: IPR000328; Env_GP41.
DR Pfam: PF00517; GP41; 1.
KW Transmembrane.
FT NON-TER 122 122
SQ SEQUENCE 122 AA; 14733 MW; 91CD821BA7A7FPCB CRC64;

Query Match
Best Local Similarity 48.4%; Score 45; DB 15; Length 122;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WIDIQGPVW 10
Db 112 WPDITQMW 121

RESULT 14
ID Q9116 PRELIMINARY; PRT: 122 AA.
AC Q9116;
DT 01-OCT-2000 (TREMBLrel, 15, Created)
DT 01-OCT-2000 (TREMBLrel, 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel, 19, Last annotation update)
DR Envelope glycoprotein (Fragment).
GN ENV.
OC Human immunodeficiency virus type 1.
OC Viruses; Retrod viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UG1636;
RA MEDLINE=20284721; PubMed=10826489;
RA Downing R., Pieniazek D., Hu D.J., Bityahaho B., Fridlund C.,
RA Rayfield M.A., Sempaia S.D., Lal R.B.;
RT "Genetic characterization and phylogenetic analysis of HIV-1 subtype C
RT in the East of Africa."
DR EMBL: AF206042; AAF81998.1; -
DR InterPro: IPR000328; Env_GP41.
DR Pfam: PF00517; GP41; 1.
KW Transmembrane.
FT NON-TER 122 122
SQ SEQUENCE 122 AA; 14710 MW; C1F846B32A3F4400 CRC64;

Query Match
Best Local Similarity 48.4%; Score 45; DB 15; Length 122;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WIDIQGPVW 10

```

```

Db 112 WPDITQMW 121

RESULT 15
ID Q90DN6 PRELIMINARY; PRT: 122 AA.
AC Q90DN6;
DT 01-DEC-2001 (TREMBLrel, 19, Created)
DT 01-DEC-2001 (TREMBLrel, 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel, 19, Last annotation update)
DR Envelope glycoprotein (Fragment).
GN ENV. Human immunodeficiency virus type 1.
OC Viruses; Retrod viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UG1636;
RA Caride E., Hertogs K., Larder B., Dehertogh P., Brindeiro R.,
RA Machado E., de Sa C.A.M., Eyre W., Passioni L.F.C., Meneses J.A.,
RA D'Aquila R.T. and others.
RT "Genotyping and phylogeny analysis of B and non-B human
RT immunodeficiency virus type 1 subtypes from patients under HAART."
RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF307695; AAL08756.1; -
DR InterPro: IPR000328; Env_GP41.
DR Pfam: PF00517; GP41; 1.
KW Transmembrane.
FT NON-TER 122 122
SQ SEQUENCE 122 AA; 14728 MW; EBF43919281BDE CRC64;

Query Match
Best Local Similarity 48.4%; Score 45; DB 15; Length 122;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WIDIQGPVW 10
Db 112 WPDITQMW 121

Search completed: March 26, 2003, 16:49:48
OOD time : 17.6577 secs

```


C/Genetics:
A/Gene: BME10279
A/Map position: 11

Query Match
Best Local Similarity 47.3%; Score 44; DB 2; Length 331;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 DIKQFWYTPFP 15
|:|:|:|:|:|:|
DB 117 WNVADYVHPTAP 129

RESULT 8
AHL469
Interleukin protein homolog 11n0295 [imported] - *Listeria innocua* (strain C11p1262)
C/Species: *Listeria innocua*
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C/Accession: AHL469
A/Status: preliminary
A/Reference number: AB1077; PMID:21517279; PMID:1167665
A/Title: Comparative genomics of *Listeria* species.
A/Description: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madheno, E.; Maitoungam, A.; Ma
O, C.; Schluerer, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,
A/Reference number: AB1077; PMID:21517279; PMID:1167665
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-361 <RIR>
A/Cross-references: GB:AL592022; PDB:1G612724; GSPDB:GN00178
A/Experimental source: strain C11p1262
C/Genetics:
A/Gene: 11n0295

Query Match
Best Local Similarity 47.3%; Score 44; DB 2; Length 361;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 IDIKQFWYTPFP 15
|:|:|:|:|:|:|
DB 310 VTLHDFWVTPFP 323

RESULT 9
T27241
Hypothetical protein Y5701C.31 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
A/Accession: T27241
A/Status: preliminary
A/Reference number: Z20330
A/Description: submitted to the EMBL Data Library, September 1997
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-425 <RIR>
A/Cross-references: EMBL:Z203281; PDB:1CML5530.1; GSPDB:GN00022; CESP:Y5701C.31
A/Experimental source: clone Y5701C
C/Genetics:
A/Gene: CESP:Y5701C.31
A/Map position: 4
A/Introns: 160/3; 187/2; 309/2; 361/3

Query Match
Best Local Similarity 47.3%; Score 44; DB 2; Length 425;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 WIDIKQFWYTPFP 15
|:|:|:|:|:|:|
DB 146 WPRIDFQWTPFP 160

RESULT 10
AB9398
COA synthetase [imported] - *Staphylococcus aureus* (strain N315)
C/Species: *Staphylococcus aureus*
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 17-May-2002
C/Accession: AB9398
A/Status: preliminary
A/Reference number: A89358
A/Title: Whole-genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A/Description: Shiba, T.; Hattori, M.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekikawa, K.;
C.; Shiba, T.; Hattori, M.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekikawa, K.;
A/Reference number: A89358
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-568 <RIR>
A/Cross-references: GB:BA000018; PDB:1G1701528; PDB:1BBA42822.1; GSPDB:GN00149
A/Experimental source: strain N315
C/Genetics:
A/Gene: acsa
C/Superfamily: acetate-CoA ligase; acetate-CoA ligase homology

Query Match
Best Local Similarity 47.3%; Score 44; DB 2; Length 568;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 DIKQFWYTPFP 14
|:|:|:|:|:|:|
DB 292 DFQYVWYTPFP 303

RESULT 11
S69206
Hypothetical protein white collar 1 - *Neurospora crassa*
C/Species: *Neurospora crassa*
C/Date: 21-Apr-1997 #sequence_revision 09-May-1997 #text_change 11-Jan-2002
C/Accession: S69206
A/Status: preliminary
A/Reference number: EMBL:J15; 1650-1657, 1996
A/Title: White collar-1, a central regulator of blue light responses in *Neurospora*, is a
A/Reference number: S69206; PMID:86203083; PMID:8612589
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1154 <BAL>
A/Cross-references: EMBL:X94300; NID:g1279576; PDB:1G1480115
C/Genetics:
A/Introns: 967/3
A/Superfamily: GATA-type zinc finger homology
A/Description: P.932-931/Domain: GATA-type zinc finger homology <GZF>

Query Match
Best Local Similarity 46.8%; Score 43.5; DB 2; Length 1154;
Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 3 DIKQFWYTPFP 14
|:|:|:|:|:|:|
DB 529 DIKQYVWYTPFP 539

RESULT 12
T14669
P-loop protein - *Yersinia pestis* plasmid pMT1
C/Species: *Yersinia pestis*
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C/Accession: T14669
A/Status: preliminary
A/Reference number: T14669
A/Title: submitted to the EMBL Data Library, March 1998
A/Description: Structural organization of virulence determinants in three *Yersinia pestis*
A/Reference number: T14669

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 16:38:14 ; Search time 29.1279 Seconds
(without alignments) 68.620 Million cell updates/sec

Title: US-10-010-667a-21
Perfect score: 93
Sequence: 1 WIDKQFWVTPPTF 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapexp 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum March 08
Maximum March 1006
Listing first 49 summaries

Database :
1. A. Geneseq 101002: *
2. /SID27/gcgdata/genseq/genseqp-emb1/AA1980.DAT: *
3. /SID27/gcgdata/genseq/genseqp-emb1/AA1981.DAT: *
4. /SID27/gcgdata/genseq/genseqp-emb1/AA1982.DAT: *
5. /SID27/gcgdata/genseq/genseqp-emb1/AA1983.DAT: *
6. /SID27/gcgdata/genseq/genseqp-emb1/AA1984.DAT: *
7. /SID27/gcgdata/genseq/genseqp-emb1/AA1985.DAT: *
8. /SID27/gcgdata/genseq/genseqp-emb1/AA1986.DAT: *
9. /SID27/gcgdata/genseq/genseqp-emb1/AA1987.DAT: *
10. /SID27/gcgdata/genseq/genseqp-emb1/AA1988.DAT: *
11. /SID27/gcgdata/genseq/genseqp-emb1/AA1989.DAT: *
12. /SID27/gcgdata/genseq/genseqp-emb1/AA1990.DAT: *
13. /SID27/gcgdata/genseq/genseqp-emb1/AA1991.DAT: *
14. /SID27/gcgdata/genseq/genseqp-emb1/AA1992.DAT: *
15. /SID27/gcgdata/genseq/genseqp-emb1/AA1993.DAT: *
16. /SID27/gcgdata/genseq/genseqp-emb1/AA1994.DAT: *
17. /SID27/gcgdata/genseq/genseqp-emb1/AA1995.DAT: *
18. /SID27/gcgdata/genseq/genseqp-emb1/AA1996.DAT: *
19. /SID27/gcgdata/genseq/genseqp-emb1/AA1997.DAT: *
20. /SID27/gcgdata/genseq/genseqp-emb1/AA1998.DAT: *
21. /SID27/gcgdata/genseq/genseqp-emb1/AA2000.DAT: *
22. /SID27/gcgdata/genseq/genseqp-emb1/AA2002.DAT: *
23. /SID27/gcgdata/genseq/genseqp-emb1/AA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score of the result being printed, and is derived by analysis of the score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	93	100.0	15	21	AAV58200
2	93	100.0	15	21	AA582788
3	93	100.0	339	21	AA158194
4	93	100.0	339	21	AA158194
5	93	100.0	339	22	AAW78845
6	93	100.0	339	22	AAW01282
7	93	100.0	339	23	ABG61813
8	93	100.0	339	23	ABG95387
9	93	100.0	374	22	AAE079829
10	93	100.0	375	22	AAE02780

11	81	87.1	339	20	AAW63109	Kidney injury asso
12	47	50.5	351	22	AAU22122	Human novel oral
13	47	50.5	351	22	AAU22122	Human novel oral
14	45	48.4	268	23	ABG68314	Envelope protein 9
15	45	48.4	269	19	AAU22837	SRD ID NO. 33 from
16	45	48.4	269	23	ABG68308	G-protein coupled
17	45	48.4	353	15	AAW48724	G-protein coupled
18	45	48.4	353	17	AAW02696	HIV-1 non-subtype
19	45	48.4	846	21	AAW69345	HIV-1 gp120 protei
20	45	48.4	861	19	AAW45074	HIV-1 gp120 protei
21	45	49.3	106	22	AAW04068	Human polypeptide
22	44	47.3	403	20	AAW97719	Scaphyloccoccus aur
23	44	47.3	568	22	AAU33996	Scaphyloccoccus aur
24	44	47.3	568	22	AAU37036	Scaphyloccoccus aur
25	44	47.3	568	22	AAU37036	Scaphyloccoccus aur
26	43	46.2	62	22	AAU43900	Protein AN70 relet
27	43	46.2	119	23	AAU72529	Protein AN70 relet
28	43	46.2	147	23	AAU72529	Protein AN70 relet
29	43	46.2	159	22	AAU72529	Protein AN70 relet
30	43	46.2	163	19	AAW69321	Protein AN70 relet
31	43	46.2	171	21	AAU77261	Protein AN70 relet
32	43	46.2	173	21	AAU77261	Protein AN70 relet
33	43	46.2	173	21	AAU77261	Protein AN70 relet
34	43	46.2	290	22	AAU56542	Protein AN70 relet
35	43	46.2	474	21	AAU77371	HIV-1 group O env
36	43	46.2	641	22	AAW61523	Protein AN70 relet
37	43	46.2	873	20	AAU77376	Protein AN70 relet
38	43	46.2	873	20	AAU77376	Protein AN70 relet
39	43	46.2	873	21	AAU77376	Protein AN70 relet
40	43	46.2	877	17	AAW07205	HIV-1 group O env
41	42	45.2	268	19	AAU22872	SRD ID NO. 66 from
42	42	45.2	268	19	AAU22872	SRD ID NO. 71 from
43	42	45.2	268	19	AAU22872	SRD ID NO. 76 from
44	42	45.2	268	19	AAU22872	SRD ID NO. 76 from
45	42	45.2	268	23	ABG69343	Envelope protein 9

ALIGNMENTS

RESULT 1
AA158200
AA158200 standard, peptide, 15 AA.
AC
AAV58200:
14-MAR-2000 (first entry)
DT
DE Human STRAP-1 peptide, corresponding to STRAP-1 extracellular region 3.
XX
XX Serpentine transmembrane action of the protease, STRAP-1, protease;
XX Serpentine transmembrane domain-type IIIa membrane protein; expression; cancer;
XX prostate cancer; bladder cancer; colon cancer; pancreatic cancer;
XX ovarian cancer; tumour antigen; immunisation; immune response;
XX cellular; humoral; anticancer vaccine; antibody/ detection; diagnosis;
XX prognosis; monitoring; susceptibility; therapeutic inhibitor;
XX drug targeting; recombinant protein.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX PN M09362941-A2.
XX
XX PD 09-DEC-1999.
XX
XX PR 01-JUN-1999, 99MO-US12157.
XX
XX PR 01-JUN-1998, 98US-0087520.
XX
XX PR 30-JUN-1998, 98US-0091183.
XX
XX (UROC-) UROGENESIS INC.
XX (AFAR/) AFAR D E.
XX (H08E/) HUBERT R S.

XX Key Location/Qualifiers
 FT 1..69 "Intracellular region 1"
 FT 70..91 "Intracellular region 1"
 FT /note="Transmembrane domain 1"
 FT 92..113
 FT /note="Extracellular region 1 (AAVS8198)"
 FT 114..136
 FT /note="Transmembrane domain 2"
 FT 137..162
 FT /note="Intracellular region 2"
 FT 163..184
 FT /note="Transmembrane domain 3"
 FT 185..218
 FT /note="Extracellular region 2 (AAVS8199)"
 FT 219..241
 FT /note="Transmembrane domain 4"
 FT 242..253
 FT /note="Intracellular region 3"
 FT 253..276
 FT /note="Transmembrane domain 5"
 FT 277..291
 FT /note="Extracellular region 3 (AAVS8200)"
 FT 292..313
 FT /note="Transmembrane domain 6"
 FT 314..339
 FT /note="Intracellular region 4"
 FT Region
 XX MO9962941.A2.
 XX 09-DEC-1999.
 XX 01-JUN-1999; 59MC-US12157.
 XX 01-JUN-1998; 98US-0087520.
 XX 30-JUN-1998; 98US-0091183.
 XX (UNCG-) UNOGENESYS INC.
 XX (AFAR/) AFAR D E.
 XX (HUBE/) HUBERT R S.
 XX (LEONG/) LEONG K.
 XX (LEONG/) LEONG K. B.
 XX (SAF/) SAFRAN D C.
 XX Afar DE, Hubert RS, Leong K, Raitano AB, Saffran DC;
 XX WPI: 2000-072832/06.
 XX N-PSDB; AA249395, AA249396.
 XX Novel proteins useful as diagnostic markers and therapeutic targets,
 XX particularly for prostatic cancer -
 XX Claim 1; Fig 1A; 83pp; English.
 XX This sequence represents a novel human protein, STRAP-1 (serpentine
 XX transmembrane antigen of the prostate). STRAP-1 is the prototype
 XX member of the STRAP family of proteins (AA158194-158197) which
 XX has a high degree of structural conservation. Previous studies
 XX have indicated that STRAP-1 is a type IIIa membrane protein and
 XX gene has been localized to chromosome 7p22. STRAP-1 is thought to be a
 XX type IIIa membrane protein and is expressed predominantly in prostate
 XX cells in normal human tissues. Structurally, STRAP-1 is a 339 amino
 XX acid protein characterized by six transmembrane domains and
 XX intracellular N- and C-terminal, suggesting that it folds in a
 XX serpentine manner into three extracellular and two intracellular loops.
 XX The protein is predicted to be a type IIIa membrane protein and
 XX throughout all stages of prostate cancer. STRAP-1 mRNA and/or protein is
 XX also overexpressed in certain other cancers, including bladder, colon,
 XX pancreatic and ovarian cancer. The function of the STRAP proteins is not
 XX known. They may be ion channels (from the presence of six transmembrane
 XX domains), a feature which is shared by certain ion channels) or
 XX gap-junction proteins (from immunohistochemical staining). STRAP-1 and

CC STRAP-2 are cell-surface tumour antigens. Immunisation with a STRAP
 CC protein induces cellular and humoral immune responses against
 CC STRAP-expressing cells. STRAP proteins may be used to identify
 CC specific antibodies against the antigen. The antibodies generate
 CC specific antibodies against the antigen. The antibodies generate
 CC and monitoring of cancers (or susceptibility to cancer), as therapeutic
 CC inhibitors or to target therapeutic agents to their site of action. STRAP
 CC nucleic acids may be used for recombinant protein production, as
 CC diagnostic and prognostic reagents, for identifying STRAP-expressing
 CC cells for screening inhibitors of STRAP expression and for therapeutic
 CC modulation/inhibition of STRAP expression. Since high levels of STRAP
 CC expression have been observed in certain human cancers, STRAP-1 may
 CC be systemically administered agents and because they are expressed mainly
 CC on prostatic epithelial cells, agents targeted to them should have
 CC minimal side effects on other tissues.
 XX Sequence 339 AA;
 XX SQ
 XX Query Match 100.0%; Score 83; DB 21; Length 339;
 XX Blast Local Similarity 100.0%; Pct Id 8.4e+06;
 XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WIDIKQFWYTPPTF 15
 DB 277 WIDIKQFWYTPPTF 291
 XX
 XX RESULT 4
 XX AAU69927
 XX ID AAU69927 standard; Protein; 339 AA.
 XX AC AAU69927;
 XX XX 30-JAN-2002 (first entry)
 XX DB Human prostate cDNA encoded protein #72.
 XX KW Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.
 XX OS Homo sapiens.
 XX PN W0200173032-A2.
 XX PS 04-OCT-2001.
 XX XX 27-MAR-2001; 2001MO-US09919.
 XX PF 27-MAR-2000; 2000US-0536857.
 XX PR 09-MAY-2000; 2000US-0568100.
 XX PR 11-MAY-2000; 2000US-0570737.
 XX PR 12-MAY-2000; 2000US-0570737.
 XX PR 27-JUN-2000; 2000US-0605783.
 XX PR 10-AUG-2000; 2000US-0656215.
 XX PR 29-AUG-2000; 2000US-0651236.
 XX PR 06-SEP-2000; 2000US-0657279.
 XX PR 02-OCT-2000; 2000US-0679426.
 XX PR 10-OCT-2000; 2000US-0685166.
 XX XX (CONTR-) CORIXA CORP.
 XX XX XU Y, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 XX PI Finger GR, Reiter MW, Stolk JA, Day CH, Vedvick TS, Carter DJ;
 XX PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
 XX DR WPI: 2001-639232/73.
 XX DR N-PSDB; AA564180.
 XX PT New human prostate-specific polypeptides and polymucleotides useful for
 XX the diagnosis and treatment of cancer, especially prostate cancer -
 XX Claim 2; Page 549; 579pp; English.
 XX The invention relates to isolated prostate-specific

CC polynucleotides, polypeptides, fusion proteins of the polypeptides, antibodies raised against the polypeptides (or antigenic epitopes derived from them) and antigen-presenting cells expressing the polypeptides, polynucleotides, polypeptides, fusion proteins of the polypeptides and antigen-presenting cells are useful for stimulating and/or expanding T cells specific for a tumour protein, and for inhibiting the development of cancer especially prostate cancer. Compositions comprising the polynucleotide and/or polypeptide are useful for stimulating an immune response, and for treating cancer. The oligonucleotide is useful for detecting cancer, and the present sequence is a prostate specific polypeptide of the invention.

SO Sequence 339 AA;

Query Match 100.0%; Score 93; DB 22; Length 339;
 Best Local Similarity 100.0%; Pred. No. 8.4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WIDIKQFWVTPPTF 15
 |||||
 Db 277 WIDIKQFWVTPPTF 291

RESULT 5
 AAM78845
 ID AAM78845 standard; Protein, 339 AA.
 AC AAM78845;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 1507.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 XX tissue growth factor; immunomodulatory; cancer; leukemia;
 KW nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 XX NO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001MO-US04098.
 XX
 PR 03-FEB-2000; 2000US-0466914.
 XX
 PR 21-NOV-2000; 2000US-0566875.
 XX
 PR 25-JUN-2000; 2000US-0620325.
 XX
 PR 19-JUL-2000; 2000US-0620325.
 XX
 PR 01-SEP-2000; 2000US-0654936.
 XX
 PR 15-SEP-2000; 2000US-0663561.
 XX
 PR 20-OCT-2000; 2000US-0693255.
 XX
 PR 30-NOV-2000; 2000US-0728422.
 XX
 XX (HSE-) HSE0 INC.
 XX
 XX
 PI Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhtman T, Goodrich R;
 XX
 XX WPI; 2001-476283/51.
 XX
 DR N-PE09; AAK51978.
 XX
 XX Nucleic acid encoding polypeptides with cytokine-like activities,
 XX useful in diagnosis and gene therapy -
 XX
 XX Claim 20; Page 3800-3801; 622pp; English.
 XX
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 XX encoded polypeptides (AAM78323-AAH0302) that exhibit activity elacting to
 XX cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides are useful in immunomodulatory activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukemia, nervous system disorders, arthritis and inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAH00020) are omitted as the relevant pages from the sequence listing were missing at the time of publication.

SO Sequence 339 AA;

Query Match 100.0%; Score 93; DB 22; Length 339;
 Best Local Similarity 100.0%; Pred. No. 8.4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WIDIKQFWVTPPTF 15
 |||||
 Db 277 WIDIKQFWVTPPTF 291

RESULT 6
 AAM01282
 ID AAM01282 standard; Protein, 339 AA.
 AC AAM01282;
 XX
 DT 04-OCT-2001 (first entry)
 XX
 DE P789P amino acid sequence.
 XX
 KW Human; prostate cancer; prostate-specific; diagnosis; vaccine; cytostatic; gene therapy; metastasis.
 XX
 OS Homo sapiens.
 XX
 XX NO200151633-A2.
 XX
 XX 19-JUL-2001.
 XX
 PF 16-JAN-2001; 2001MO-US01574.
 XX
 PR 14-JAN-2000; 2000US-0483672.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
 XX Krole MD, Renger GR, Day CH, Reuter WM, Seok JA, Skelky VM;
 PI Wang A, Weagner MJ;
 XX
 XX WPI; 2001-425873/45.
 XX
 DR New polynucleotide encoding a prostate-specific protein, for
 XX diagnosing, monitoring and treating prostate cancer in a patient and
 XX for use in vaccines -
 XX
 XX Claim 2; Page 510-512; 543pp; English.
 XX
 XX The present invention describes polynucleotide sequences (I) which encode
 XX prostate-specific proteins (II). (I) and (II) have cytostatic activity,
 XX and can be used in vaccine production and gene therapy. (I), (II),
 XX antibodies to (II), fusion proteins comprising (I), and isolated
 XX cells expressing (I) are useful in the diagnosis and treatment of
 XX (I) and the antibodies are also used in the detection of cancer in a
 XX patient. The cancer that is diagnosed or treated is particularly
 XX prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
 XX (I) can be used for monitoring the progression of cancer in a patient.
 XX (I) and (II) can also be used to improve diagnostic and therapeutic
 XX methods for prostate cancer. They can indicate the level of metastasis
 XX as well as the prostate volume. AAK5357 to AAK5394 and AAM0115 to
 XX AAM0138 represent polynucleotide and amino acid sequences used in the

XX MO200140276-A2.
 XX 07-JUN-2001.
 XX 06-DEC-2000; 2000MO-US33040.
 XX 06-DEC-1999; 99US-0455486.
 XX (URQC-) UROGENESIS INC.
 XX After DRH, Hubert RS, Raicano AB, Saffran DC, Mitchell SC, Faris M,
 XX Jakobovits A.
 XX NPI; 2001-367804/38.
 XX N-PSDB; AAD07067.
 XX New STEAP (six transmembrane epithelial antigen of the prostate)
 XX protein, expressed in human cancer, useful for detecting and treating
 XX cancer.
 XX Example 2; Fig 1A-1B; 187dp; English.
 XX The present sequence is human six transmembrane epithelial antigen of
 XX the prostate (STEAP)-1 protein of clone 10. STEAP-1 gene is located on
 XX surface serpentine transmembrane antigens. STEAP-1 gene is located on
 XX chromosome p22 and is used in gene therapy. Including the development
 XX of a cancer vaccine, STEAP-1 gene is used in gene therapy. STEAP-1
 XX and pancreatic) expressing STEAP or inhibiting growth or killing cells
 XX expressing STEAP in a patient, comprises administering a vaccine
 XX composition to the patient. Treating a patient with a cancer that
 XX expresses STEAP, or inhibiting growth or killing cells expressing STEAP,
 XX comprises administering to the patient a vector encoding single chain
 XX monoclonal antibody that comprises the variable domain of the heavy and
 XX light chain of the antibody. The vector encodes a single chain antibody
 XX such that the vector delivers the single chain monoclonal antibody coding
 XX sequence to the cancer cells and the encoded single chain monoclonal
 XX antibody is expressed intracellularly.
 XX Note: The present sequence is also shown in sequence listing of the
 XX specification, but it lacks amino acid residues at its N-terminal end.
 XX
 XX Sequence 375 AA;
 XX
 XX Query Match 100.0%; Score 93; DB 22; Length 375;
 XX Best Local Similarity 100.0%; Pred. No. 9,2e-06;
 XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 WIDIKQFWWTPPTF 15
 XX I:|||||
 XX 277 WIDIKQFWWTPPTF 291
 XX
 XX RESULT 11
 XX AAM6309
 XX ID AAM6309 standard; Protein; 339 AA.
 XX AC AAM6309;
 XX XX
 XX 01-MAR-1999 (first entry)
 XX XX
 XX Kidney injury associated molecule HM018 protein.
 XX XX
 XX Kidney injury associated molecule; kidney injury related molecule;
 XX KIM; tissue growth promotion; regeneration; renal condition;
 XX acute renal failure; acute nephritis; tumour.
 XX Ractua sp.
 XX MO965071-A1.
 XX 26-NOV-1998.
 XX PD
 XX 22-MAY-1998; 98MO-US10547.

XX 23-MAY-1997; 97US-0047491.
 XX 23-MAY-1997; 97US-0047490.
 XX (BIO) BIOGEN INC.
 XX Cate RL, Hession CA, Sanicola-Nadel M, Wei H;
 XX NPI; 1999-045312/04.
 XX N-PSDB; AAM60586.
 XX Kidney injury-associated molecule, KIM, polypeptides - upregulated
 XX in injured or regenerating tissues, useful to promote tissue growth
 XX and regeneration, especially to treat renal conditions
 XX Claim 17; Page 57-58; 21dp; English.
 XX The present sequence represents a kidney injury associated molecule
 XX protein. The protein is a polypeptide chain encoded by a gene
 XX by the protein KIM encoding chain. The protein promotes growth and/or
 XX survival of damaged tissue (e.g. renal tissue), since the KIM proteins
 XX are upregulated in injured or regenerating (especially renal) tissues.
 XX KIM fusion proteins, conjugates, antibodies and vectors can also be used
 XX therapeutically, e.g. these or the KIM proteins may be included with an
 XX acceptable carrier in pharmaceutical compositions, useful for therapy/
 XX prophylaxis of conditions associated with dysfunction/dysregulation of
 XX the kidney. The protein is expected to be useful in the treatment of renal
 XX function in patients with renal diseases (nephritis). The
 XX polypeptides can be used to produce antisense sequences which, when
 XX internalised into cells, can disrupt expression of a cellular KIM gene,
 XX also useful in therapy (e.g. to block the growth of tumours dependent on
 XX KIM for growth) or compositions. The proteins and polynucleotides are
 XX useful diagnostically e.g. to detect and quantify renal injury/disease
 XX indicative of increased risk, or presence of, renal injury or impaired
 XX renal function. The proteins and polynucleotides are also useful in
 XX increased risk, or presence of an autoimmune response or abnormal
 XX tissue growth arising from/affecting renal tissue). The proteins can
 XX also be used to locate KIM-producing cells (especially specific loci,
 XX e.g. tissue masses abnormally producing/expressing KIM such as tumours
 XX arising from/affecting renal tissue), by contacting cells with an
 XX imageable KIM-binding reagent and imaging reagent accumulation.
 XX
 XX Sequence 339 AA;
 XX
 XX Query Match 87.1%; Score 81; DB 20; Length 339;
 XX Best Local Similarity 80.0%; Pred. No. 0.00047;
 XX Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 XX
 XX QY 1 WIDIKQFWWTPPTF 15
 XX I:|||||
 XX 277 WIDIKQFWWTPPTF 291
 XX
 XX RESULT 12
 XX AAU2122
 XX ID AAU2122 standard; Protein; 51 AA.
 XX AC AAU2122;
 XX XX
 XX 17-DEC-2001 (first entry)
 XX XX
 XX Human novel foetal antigen, SEQ ID NO 1456.
 XX XX
 XX Human, foetal tissue antigen; antiinflammatory; neuroprotective;
 XX immunomodulator; cardiovascular; cytoprotective; nephroprotective;
 XX cardiovascular; autoimmune disease; rheumatoid arthritis;
 XX cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
 XX cerebral ischaemia; angiogenesis; nervous system disorder;
 XX Alzheimer's disease; infection; ocular disorder; corneal infection;
 XX wound healing; epithelial cell proliferation; food additive.
 XX Homo sapiens.
 XX OS

XX WC0200155312-A2.
XX 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01331.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 16-FEB-2000; 2000US-0185458.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0196076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224617.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228932.
PR 01-SEP-2000; 2000US-0228287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0230513.
PR 05-SEP-2000; 2000US-0230518.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 11-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0233401.
PR 14-SEP-2000; 2000US-0233402.
PR 14-SEP-2000; 2000US-0233403.
PR 14-SEP-2000; 2000US-0233405.
PR 14-SEP-2000; 2000US-0233406.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
XX
PR 27-SEP-2000; 2000US-0235836.
PR 28-SEP-2000; 2000US-0236327.
PR 28-SEP-2000; 2000US-0236328.
PR 28-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241121.
PR 20-OCT-2000; 2000US-0241121.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0246517.
PR 08-NOV-2000; 2000US-0246514.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246529.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251486.
PR 06-DEC-2000; 2000US-0251487.
PR 06-DEC-2000; 2000US-0251489.
PR 06-DEC-2000; 2000US-0251489.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HDMV -) HUMAN GENOME SCI INC.
XX

XX Bristeed AC, Judice JK, McDowell RS, Phealan JC, Starovansnik MA,
 PI Wells JA;
 XX WPI, 2002-487624/52.
 XX New cyclic peptides from human immune deficiency virus gp41, useful for
 PT treatment or prevention of HIV infection, are constrained to have
 PT alpha-helical conformation -
 PS Disclosure: Column 189-192; 175pp; English.
 XX The invention relates to cyclic peptides (A) with a constrained helical
 CC conformation, derived from gp41 (glycoprotein 41, a viral envelope
 CC protein) of human immunodeficiency virus (HIV). The cyclic
 CC peptides have formulas given in the specification part of which are
 CC derived from a consensus sequence of gp41 derived from HIV clades A, B,
 CC C, D, E or O. The peptides are used to cause induction of a specific
 CC immune response, resulting in antibodies that prevent infection or at risk
 CC of HIV infection, either as antihuman/anti-infection agents or, at risk
 CC of HIV infection, to raise antibodies. The antibodies may be used for diagnosis or
 CC vaccine) to raise antibodies. The antibodies may be used for diagnosis or
 CC syndrome, AIDS), e.g. prevention of mother-to-child transmission or in
 CC cases of health care accidents. The peptides can be based on specific HIV
 CC strains, e.g. gp41, gp120, gp160, gp170, gp180, gp190, gp200, gp210, gp220,
 CC or gp230, or a combination of them should cover a wide range of
 CC protection. The present sequence is gp41 protein from a particular
 CC HIV clade used to derive a consensus sequence of gp41.
 XX SQ Sequence 268 AA;
 XX
 XX Query Match 48 4%; Score 45; DB 23; Length 268;
 XX Best Local Similarity 60.0%; Pred. No. 69;
 XX Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WIDIKQPVWY 10
 DB 203 WPDITQWLMT 211
 XX
 XX RESULT 15
 XX AAAY22837
 XX AAAY22837 standard; Protein; 269 AA.
 XX AC AAAY22837;
 XX DT 19-AUG-1999 (first entry)
 XX DT 19-AUG-1999 (first entry)
 XX XX SEQ ID NO. 33 from MO9820036.
 XX XX
 XX XX HIV, gp41 protein; constrained helical peptide; HIV infection;
 XX vaccine; antibody; viral membrane fusion; viral infectivity;
 XX 119and affinity purification; protein A replacement;
 XX immunoglobulin purification; epitope mimic.
 XX XX Human immunodeficiency virus.
 XX XX MO9820036-A1.
 XX PD 14-MAY-1998.
 XX PF 05-NOV-1997; 97MO-US20069.
 XX XX 16-JUN-1997; 97US-0876598.
 XX PR 06-NOV-1996; 96US-0743698.
 XX XX
 XX PA (GETH) GENENTECH INC.
 XX XX Bristeed A, Judice JK, McDowell RS, Phealan JC, Starovansnik MA,
 XX PI Wells JA;
 XX PI

DR WPI, 1998-286866/25.
 XX Production of constrained helical peptide(s) by linking side chains
 XX on termini of octa-peptide - derived from human immunodeficiency
 PT virus gp41 protein, useful in vaccines for treatment and prevention
 PT of infection
 XX Claim 11; Page 174-175; 275pp; English.
 XX Peptides AAAY22805-Y22917 are derived from human immunodeficiency virus
 XX (HIV). Specifically AAAY22810-Y22910 are derived from gp41 proteins
 CC of known HIV virus strains (AAAY22810, AAAY22871, AAAY22880, AAAY22888 and
 CC AAAY22903 represent consensus sequences of various sections of the gp41
 CC protein). Sequences derived from the peptides are used to produce
 CC constrained helical peptides of the invention. The constrained helical
 CC peptide is produced by synthesizing an octapeptide in which both terminal
 CC amino acids have a side-chain that includes a group substituted side-chain
 CC bond, and C-terminal amino acids are linked to produce two amide bonds
 CC The constrained helical peptides are used to create or prevent HIV
 CC infection, especially as vaccines that generate antibodies that
 CC prevent viral membrane fusion or infectivity. Vaccines may contain
 CC constrained helical peptides derived from several different strains of
 CC HIV. The antibodies are also useful for diagnosing HIV infection. Other
 CC uses for the constrained helical peptides include: (a) as a binding protein
 CC available as a replacement for protein A in immunoglobulin
 CC purification); as epitope mimics for antibody production; for isolation
 CC of synthetic antibody clones from phage display libraries, or as stable
 CC forms of "floppy" peptides or proteins.
 XX SQ Sequence 269 AA;
 XX
 XX Query Match 48 4%; Score 45; DB 19; Length 269;
 XX Best Local Similarity 60.0%; Pred. No. 69;
 XX Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WIDIKQPVWY 10
 DB 203 WPDITQWLMT 212
 XX
 XX Search completed: March 26, 2003, 16:48:16
 XX Job time : 30.1279 secs

CC	-1- SUBUNIT: HETEROODIMER OF TWO SUBUNITS, ONE OF WHICH BINDS GTP.
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (probable).
CC	-1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC	EF2 SUBFAMILY.
CC	-----
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL consortium -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to ef2@isb-sib.ch).
CC	-----
DR	EMBL, M21129, AA33133.1, -
DR	EMBL, X07163, CA30155.1, -
DR	EMBL, Y00829, CA36760.1, -
DR	EMBL, Z46727, CA36677.1, -
DR	PIR, S00733, EFEBY52.
DR	SCD, S0020579, S0P5.
DR	InterPro, IPR004160, EFMD2.
DR	InterPro, IPR000795, EF_GTPbind.
DR	Pfam, PF00069, GTP_EFTU.1.
DR	Pfam, PF03143, GTP_EFTU_D3.1.
DR	Pfam, PF03144, GTP_EFTU_D2.1.
DR	ProSITE, PS00301, EFCACTOR_GTP.1.
KW	Protein biosynthesis; Repeat; Phosphorylation; GTP-binding.
FT	DOMAIN 135..148
FT	CDIPTN CHARGED SORT OF REPEATS.
FT	DOMAIN 135..148
FT	SIMILAR 254..655
FT	NP_BIND 267..274
FT	NP_BIND 344..348
FT	NP_BIND 406..409
FT	ACT_SITE 273..273
FT	ACT_SITE 407..407
FT	ACT_SITE 437..437
FT	CONFLICT 53..53
FT	CONFLICT 53..53
SO	SEQUENCE 685 AA; 76551 MW; 43912AD07DF6A15 CRC64;
QY	2 IDIKOPHPTPT 14
Db	459 VDPKCPVITGFT 471
RESULT 8	
TLMT MOUSE	STANDARD. PRT; 795 AA.
ID	TUATL1_MOUSE
CC	OS=TOULOUSE
CC	DB=TOULOUSE
CC	DT=15-JUN-2002 (Ref. 4). Created
CC	DT=15-JUN-2002 (Ref. 4). Last annotation update
CC	DE TOLL-like receptor 1 precursor (Toll/interleukin-1 receptor-like)
CC	(TIR).
CC	OS Mus musculus (Mouse)
CC	NCBI_TaxID=10090
CC	NCBI_TaxID=10090
CC	SEQUENCE FROM N. A.
CC	STRAIN=BALE/c; TISSUE=Macrophage;
CC	MEDLINE=20555851; PubMed=11095740.
CC	OSMayer, A., Underhill, D.M., Fomenko, J.D., Hajjar, A.M., Smith K.D.,
CC	"The repertoire for pattern recognition of pathogens by the innate
CC	immune system is defined by cooperation between TOLL-like
CC	receptors",
CC	Proc. Natl. Acad. Sci. U.S.A. 97:13766-13771(2000).
CC	[2]

[illegible]

FT	CAROHYD	239	239	N-LINKED (GLCNAc..)	(POTENTIAL..)			
FT	CAROHYD	250	260	N-LINKED (GLCNAc..)	(POTENTIAL..)			
FT	CAROHYD	274	274	N-LINKED (GLCNAc..)	(POTENTIAL..)			
FT	CAROHYD	293	293	N-LINKED (GLCNAc..)	(POTENTIAL..)			
FT	CAROHYD	328	328	N-LINKED (GLCNAc..)	(POTENTIAL..)			
FT	CAROHYD	336	336	N-LINKED (GLCNAc..)	(POTENTIAL..)			
FT	CAROHYD	352	352	N-LINKED (GLCNAc..)	(POTENTIAL..)			
FT	CAROHYD	382	382	N-LINKED (GLCNAc..)	(POTENTIAL..)			
FT	CAROHYD	388	388	N-LINKED (GLCNAc..)	(POTENTIAL..)			
FT	CAROHYD	392	392	N-LINKED (GLCNAc..)	(POTENTIAL..)			
FT	CAROHYD	401	401	N-LINKED (GLCNAc..)	(POTENTIAL..)			
FT	CAROHYD	438	438	N-LINKED (GLCNAc..)	(POTENTIAL..)			
FT	CAROHYD	454	454	N-LINKED (GLCNAc..)	(POTENTIAL..)			
FT	CAROHYD	602	602	N-LINKED (GLCNAc..)	(POTENTIAL..)			
FT	CAROHYD	607	607	N-LINKED (GLCNAc..)	(POTENTIAL..)			
FT	CAROHYD	616	616	N-LINKED (GLCNAc..)	(POTENTIAL..)			
FT	CAROHYD	628	628	N-LINKED (GLCNAc..)	(POTENTIAL..)			
SO	SEQUENCE	847 AA,	96335 MW,	DA901317BD/PF2AB	CRC64			
Query Match		45.2%	Score 42,	DB 1,	Length 847;			
Best local similarity		50.0%	Prod. No. 35,					
Matches		5;	Conservative	3;	Mismatches 2;			
					Indels 0;			
					Gaps 0;			
Qy	1	WIDINQPVWY	10					
Db	663	WFDISNMW	672					
RESULT 10								
Y89	YEAST	STANDARD;	PRT	224	AA.			
ID	Y899	YEAST						
AC	P5321;							
DT	01-OCT-1996 (Ref. 34, Created)							
DT	01-OCT-1996 (Ref. 34, Last sequence update)							
DT	01-OCT-1996 (Ref. 34, Last annotation update)							
GN	Hypothetical 25.3 kDa protein in TIM23-ARE2 intergenic region.							
DB	Y89018W OR N3185.							
OC	Saccharomycetes; Ascomycota (Baker's Yeast).							
OC	Euclaryota; Fungi; Eumycota; Saccharomycotina; Saccharomycetes;							
OC	Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.							
PK	11;_taxid=4932;							
RP	SEQUENCE FROM N.A.							
AP	P01 T.M.;							
RL	Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.							
CC	-1- SUBCELLULAR LOCATION: integral membrane protein (Potential).							
CC	This SWISS-PROT entry is copyrighted by the Swiss Institute of Bioinformatics.							
CC	Use by non-profit institutions is allowed. The copyright owner does not allow redistribution of this entry without prior written permission.							
CC	Use by non-profit institutions as long as its content is in no way misrepresented and the EMBL copyright notice is included.							
CC	modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/).							
CC	or send an email to license@sib-sib.ch).							
DR	EMBL; GenBank; DDBJ							
DR	006931	006931	11					
DR	006931	Y89018W						
KW	Hypothetical protein; Transmembrane.							
FT	TRANSMEM	19	39	POTENTIAL.				
FT	TRANSMEM	50	70	POTENTIAL.				
FT	TRANSMEM	152	172	POTENTIAL.				
FT	DOMAIN	196	199	POLY-GUT.				
SO	SEQUENCE	224 AA;	25344 MW;	FAZC528A008CFRC	CRC64;			
Query Match		44.6%	Score 41.5;	DB 1;	Length 224;			
Best local similarity		57.4%	Prod. No. 11;					
Matches		8;	Conservative	1;	Mismatches 4;			
					Indels 1;			
					Gaps 1;			
Qy	1	WIDINQPVWY	14					

FT TRANSMEM 209 229 POTENTIAL.
 FT TRANSMEM 233 253 POTENTIAL.
 FT TRANSMEM 274 294 POTENTIAL.
 FT TRANSMEM 310 330 POTENTIAL.
 FT TRANSMEM 369 389 POTENTIAL.
 FT TRANSMEM 398 418 POTENTIAL.
 SQ SEQUENCE 430 AA; 44661 MW; 9FDD18D100A0572E CRC64;
 Query Match 44.1%; Score 41; DB 1; Length 430;
 Best Local Similarity 33.3%; Pred. No. 26;
 Matches 5; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 QY 1 WIDIOQPMVPTPT 15
 I : : : : :
 Db 215 WLVHPSLPCPGPTT 229

RESULT 13
 ID ENV_HV1Y2 STANDARD; PRT; 703 AA.
 AC P1437;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Arylphorin beta subunit precursor.
 OS Manduca sexta (tobacco hawkmoth) (tobacco hornworm).
 NC Insecta: Lepidoptera; Manduca sexta; Lepidoptera; Glossata;
 NC Insecta: Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
 OC Diptera; Sphingidae; Sphingidae; Sphingidae; Manduca.
 NX NCBI_TaxId=7130;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA TISSUE=Larval fat body;
 RA PubMed=2898410;
 RA MIM=603702; PubMed=2898410;
 RT WU; Li, H., Burgess C.J., Price R.W., Hahn B.H.,
 RT "DNA and gene sequence of Manduca sexta arylphorin, an aromatic
 RT amino acid-rich larval serum protein. Homology to arthropod
 RT haemocyanins.";
 RL J. Biol. Chem. 264:19052-19053(1989).
 CC -1- FUNCTION: ARYLPHORIN IS A LARVAL STORAGE PROTEIN (LSP) WHICH MAY
 CC SERVE AS A STORAGE PROTEIN USED PRIMARILY AS A SOURCE OF AROMATIC
 CC AMINO ACIDS FOR THE SYNTHESIS OF THE CUTICLE, AND SERVES
 CC AS A CARRIER FOR ECDYSTEROID HORMONE.
 CC -1- SUBUNIT: ARYLPHORIN IS A HEAVYER OF SUBUNITS ALPHA AND BETA.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: FAT BODY.
 CC -1- SIMILARITY: TO ARMB, TO B-MOR STORAGE PROTEINS 1 AND 2, AND TO
 CC ARTHRODIN HEMOCYANINS.
 CC -1- ARTHRODIN HEMOCYANINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL; M28195; AAA23104.1;
 CC EMBL; M28197; AAA23105.1;
 CC PIR; B34434; B34434.
 DR HSBP; P04253; 10XV.
 DR InterPro; IP000896; Hemocyanin.
 DR InterPro; IP000540; hemocyanin.C
 DR InterPro; IP000540; hemocyanin.N
 DR Pfam; PF01322; hemocyanin.N.1.
 DR Pfam; PF01323; hemocyanin.C.1.
 DR PRINTS; PR00187; HAEMOCYANIN.
 DR PROSITE; PS00209; HAEMOCYANIN 1; 1.
 DR PROSITE; PS00210; HAEMOCYANIN 2; 1.
 KW Signal; Storage protein; Glycoprotein; Multigene family.

FT SIGNAL 1 16 ARYLPHORIN BETA SUBUNIT.
 FT CHAIN 17 703 N-LINKED (GLCNAc...) (POTENTIAL).
 FT CARBOHYD 72 712 N-LINKED (GLCNAc...) (POTENTIAL).
 FT CARBOHYD 211 212 N-LINKED (GLCNAc...) (POTENTIAL).
 SQ SEQUENCE 703 AA; 83848 MW; 2487DBB80023FA CRC64;
 Query Match 44.1%; Score 41; DB 1; Length 703;
 Best Local Similarity 60.0%; Pred. No. 42;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 3 DIQOPVPTPT 12
 I : : : : :
 Db 287 DIPESWSP 296

RESULT 14
 ID ENV_HV1Y2 STANDARD; PRT; 843 AA.
 AC P15961;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope glycoprotein GP150 precursor (Concatal; Exterior membrane
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP11)).
 GN ENV.
 OS Human immunodeficiency virus type 1 (VU-2 isolate) (HIV-1).
 NC Virusae; Retrod virusae; Retroviridae; Lentivirus.
 NC Virusae; Retrod virusae; Retroviridae; Lentivirus.
 OC Virusae; Retrod virusae; Retroviridae; Lentivirus.
 NX NCBI_TaxId=35377;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA MEDLINE=93021387; PubMed=1404605;
 RA Li, Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
 RA Shaw G.M.,
 RA "Complete nucleotide sequence, genome organization, and biological
 RA properties of HIV-1 in vivo: evidence
 RA for a limited defectiveness and complementation.";
 RL J. Virol. 66:6587-6600(1992).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL; M93258; -; NOT ANNOTATED CDS.
 CC PIR; H44001; H44001.
 DR InterPro; IP0000129; Env GP41.
 DR InterPro; IP000517; GP120.
 DR Pfam; PF00517; GP41; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Glycoprotein; Transmembrane;
 KW Signal.
 FT SIGNAL 1 29
 FT CHAIN 30 489 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT TRANSMEM 738 755 TRANSMEMBRANE GLYCOPROTEIN.
 FT TRANSMEM 738 755 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 118 201 BY SIMILARITY.
 FT DISULFID 125 192 BY SIMILARITY.
 FT DISULFID 130 155 BY SIMILARITY.
 FT DISULFID 214 243 BY SIMILARITY.
 FT DISULFID 224 235 BY SIMILARITY.
 FT DISULFID 232 326 BY SIMILARITY.
 FT DISULFID 326 345 BY SIMILARITY.
 FT DISULFID 345 405 BY SIMILARITY.
 FT CARBOHYD 87 87 N-LINKED (GLCNAc...) (POTENTIAL).
 FT CARBOHYD 129 129 N-LINKED (GLCNAc...) (POTENTIAL).
 FT CARBOHYD 135 135 N-LINKED (GLCNAc...) (POTENTIAL).
 FT CARBOHYD 138 138 N-LINKED (GLCNAc...) (POTENTIAL).
 FT CARBOHYD 154 154 N-LINKED (GLCNAc...) (POTENTIAL).
 FT CARBOHYD 158 158 N-LINKED (GLCNAc...) (POTENTIAL).

Genome version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 16:38:14 ; Search time 9.88372 Seconds
(without alignments)
142.679 Million cell updates/sec

Title: US-10-010-667a-20

Sequence: 1 RRSRYKLLMAYVQVQNKEDMATEHDVWMEI 34

Scoring table: BLASTING62
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	192	100.0	339	1 STEA_HUMAN	G94688 homo sapien
2	58.5	30.5	1418	1 CENL1_MOUSE	P34644 Caenorhabdi
3	58.5	30.5	1418	1 CENL1_MOUSE	P34644 Caenorhabdi
4	58.5	30.5	1418	1 CENL1_MOUSE	P34644 Caenorhabdi
5	52	27.1	1813	1 UN13_MOUSE	P54566 bacillus su
6	52	27.1	1813	1 UN13_MOUSE	P54566 bacillus su
7	51	26.6	495	1 GLPK_SYNY3	P27715 caenorhabd
8	51	26.6	495	1 GLPK_SYNY3	P27715 caenorhabd
9	50	26.0	784	1 ALP4_SCHRO	P74460 synchocyst
10	50	26.0	784	1 ALP4_SCHRO	P74460 synchocyst
11	49	26.0	5430	1 SV13_RAT	O92910 ratuon nov
12	49	26.0	5430	1 SV13_RAT	O92910 ratuon nov
13	49	26.0	5430	1 SV13_RAT	O92910 ratuon nov
14	49.5	25.8	350	1 DAPD_BURCI	O94328 bacillus su
15	49.5	25.8	350	1 DAPD_BURCI	O94328 bacillus su
16	49.5	25.8	350	1 DAPD_BURCI	O94328 bacillus su
17	49	25.5	512	1 YAMA_HISIN	O92987 chlamydia p
18	49	25.5	512	1 YAMA_HISIN	O92987 chlamydia p
19	49	25.5	512	1 YAMA_HISIN	O92987 chlamydia p
20	48	25.3	339	1 NCM2_HUMAN	P55579 thiazolum s
21	48.5	25.3	339	1 NCM2_HUMAN	P55579 thiazolum s
22	48.5	25.3	339	1 NCM2_HUMAN	P55579 thiazolum s
23	48	25.0	864	1 AGUL_MOUSE	O33432 mus musculu
24	48	25.0	864	1 AGUL_MOUSE	O33432 mus musculu
25	48	25.0	864	1 AGUL_MOUSE	O33432 mus musculu
26	47.5	24.7	942	1 ENVI_CABVE	P31627 caprine arc
27	47.5	24.7	942	1 ENVI_CABVE	P31627 caprine arc
28	47.5	24.7	942	1 ENVI_CABVE	P31627 caprine arc
29	47.5	24.7	942	1 ENVI_CABVE	P31627 caprine arc
30	47	24.5	244	1 CYBH_BRAJA	P27175 glucosylact
31	47	24.5	244	1 CYBH_BRAJA	P27175 glucosylact
32	47	24.5	244	1 CYBH_BRAJA	P27175 glucosylact
33	47	24.5	244	1 CYBH_BRAJA	P27175 glucosylact

34	47	24.5	1133	ATX3_TETTR	O95050 tetrathymena	
35	47	24.5	1133	ATX3_TETTR	O95050 tetrathymena	
36	47	24.5	1133	ATX3_TETTR	O95050 tetrathymena	
37	46.5	24.2	215	CIB2_RAT	O60939 homo sapien	
38	46.5	24.2	463	SYN_CLOAB	P54900 rctueu novy	
39	46.5	24.2	508	O166_BACSC	P29093 bacillus sp	
40	46.5	24.2	555	O166_BACSC	P29093 bacillus sp	
41	46	24.0	220	D1B7_MOUSE	O94510 bacillus co	
42	46	24.0	220	D1B7_MOUSE	O94510 bacillus co	
43	46	24.0	226	B1R6_HUMAN	O94911 gorilla gor	
44	46	24.0	236	B1R6_HUMAN	O94911 gorilla gor	
45	46	24.0	236	B1R6_HUMAN	O94911 gorilla gor	
46	24.0	242	1	D1B6_MOUSE	O95462 pan troglod	
45	46	24.0	242	1	D1B6_MOUSE	mus musculu

ALIGNMENTS

RESULT 1	STEA_HUMAN	STANDARD	PRT	339 AA.
AC	G94688	095034		
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DE	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Six transmembrane epithelial antigen of prostate.			
GN	STEAP OR STEAP2			
GN	STEAP OR STEAP2 (human)			
OS	Bakeryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
NCBI	TaxID=9606;			
LN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20056271; PubMed=10586738; Rastegar S., Leong K.,			
RA	Michael R.S., Vaidyanathan R., Zhou C., Raitano A.B.,			
RA	Jakobovits A., Saffran D.C., Afari D.E.H.;			
RT	"STEAP: a prostate-specific cell-surface antigen highly expressed in			
RT	human prostate tumors." J.			
RL	Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528(1999).			
LN	[2]			
RP	SEQUENCE FROM N.A.			
RX	EMBL=AF166249; competing T., Langdon Y., Maupin R.;			
RA	Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.			
LN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Skin;			
RA	Strauberg R.;			
RA	Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.			
CC	-1- SHIBLUDIAN LOCATION: Integral membrane protein (potential).			
CC	-1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PROSTATE TUMORS.			
CC	-1- This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-ch).			
DR	EMBL; AF166249; AAF17479.1; -			
DR	EMBL; AC005053; AAC79150.1; ALT INIT.			
DR	EMBL; AC004969; AAD15620.1; ALT INIT.			
DR	EMBL; BC011802; AAH11802.1; -			
DR	Gene; HGNC:11379; STEAP.			
DR	MTM; 604415; - Antigen.			
FT	TRANSMEM 71 91.			POTENTIAL.
FT	TRANSMEM 119 139			POTENTIAL.
FT	TRANSMEM 164 184			POTENTIAL.
FT	TRANSMEM 218 238			POTENTIAL.
FT	TRANSMEM 258 278			POTENTIAL.
FT	TRANSMEM 339 359			POTENTIAL.
SO	SEQUENCE 339 AA; 39851 MW; 35442A10C870187 CRC64;			

[illegible]

RC STRAIN-B1.e1c1.n2;
RX MEDLINE-91286538; PubMed-2062851;
RA Maruyama I.N., Bremner S.;
RT The phenol esterase-like unc-13
RN protein encodes a novel calcium-binding protein.
J Biol. Chem. 273(19):11392-11397 (1998).
PMOC. Nucleic Acids Res. 26(19):5729-5733 (1998).
[3]
RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
R STRAIN-B1.e1c1.n2;
RX MEDLINE-94037755; PubMed-11029047;
RA Kohm R.E., Duert J.S., Mechanic J.R., Duke A., Rakov T.L., Maruyama H.,
RT Expression of multiple UNC-13 proteins in the Ctenocephalides felis
RN nervous system".
Mol. Biol. Cell 11:3441-3452(2000).
[3]
RC SEQUENCE FROM N.A. (ISOFORMS 1 AND 4).
R STRAIN-B1.e1c1.n2;
RX Girdler A.E., Lloyd C.;
RT Submitted (MAR-1997) to the EMBL/Genbank/DDbj databases.
RN [4]
ZINC AND PHORBOL-ESTER BINDING.
MEDLINE-93075060; PubMed-1445555;
RA Ahmed S., Maruyama I.N., Komra R., Lee J., Bremner S., Lim L.;
RT "The Ctenocephalitis felis unc-13 gene product is a phospholipid-
dependent high-affinity phorbol ester receptor".
Biochem. J. 287(1995-1999)(1997).
RL - TRANSDUCING THE SIGNAL FROM DIACYLGlycerol TO EFFECTOR
FUNCTIONS. ONE SUCH FUNCTION COULD BE THE RELEASE OF
NEUROTRANSMITTER FROM NEURONS.
CC - ALTERNATIVE PRODUCTS: 4 Isoforms; 1/ZK534.2a (shown here), 2, 3
- DOMAIN: THE PHORBOL ESTER BINDING ACTIVITY IS ZINC AND CALCIUM-
DEPENDENT.
CC - DIVERGENT.
CC - DIVERGENT ANALOGOUS NOTATIONS IN UNC-13 CAUSE DIVERSE NERVOUS SYSTEM
EFFECTS.
CC - SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
BINDING DOMAIN.
CC - SIMILARITY: CONTAINS 3 CC DOMAINS.
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation at
Cambridge. All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
electronic, mechanical, photocopying, recording, or otherwise, without prior
written permission from the copyright holder. This entry is made available
under a non-exclusive license for use by non-profit organizations as long as its
content is in no way misrepresented and provided the original source is
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-ebi.ch/announce/>)
or send an email to license@isb-ebi.ch).
EMBL; ME28310; AAA93094.1; -
EMBL; US0713; AAA93780.1; -
EMBL; US0713; AAA93780.1; -
EMBL; Z297739; CAB07173.1; -
EMBL; Z79694; CAB07173.1; JOINED.
EMBL; Z73912; CAB07173.1; JOINED.
EMBL; Z73912; CAB98146.1; -
EMBL; Z73912; CA98147.1; -
EMBL; Z79694; CA98147.1; JOINED.
EMBL; Z297739; CA98147.1; JOINED.
EMBL; Z297739; CA98147.1; JOINED.
EMBL; Z297739; CAB01966.1; JOINED.
EMBL; Z73912; CAB01966.1; JOINED.
PIR; A41101; A41101.
HSP; P28667; LPFO.
DR Wormpep; ZK534.2a; CE15371.
DR Wormpep; ZK534.2b; CE15287.
INTERPRO: IPR0001008; CA; CE15287.
DR InterPro; IPR0001008; CA; CE15287.
DR Pfam; PF00110; DALD; PE-Bind; 1.
DR Pfam; PF00168; CZ; 3.
DR PRINTS; PR00360; CCDOMAIN.
DR PRINTS; PR00008; DAGPEPDOMAIN.
DR SMART; SM00109; CI; 1.

CC	DR	SMART, SMO0239; C2; 3.
CC	DR	PROSITE, PSS0049; C2; DOMAIN 1; 1.
CC	DR	PROSITE, PSS0049; C2; DOMAIN 1; 1.
CC	DR	PROSITE, PSS0081; DAG DE BIND DOM_2; 1.
CC	KW	Photo-l-ester binding; Zinc; Repeat; Alternative splicing.
CC	FT	DOMAIN 21 96
CC	FT	DOMAIN 693 742
CC	FT	DOMAIN 802 908
CC	FT	DOMAIN 1633 1738
CC	FT	VARSPLIC 1 85
CC	FT	VARSPLIC 1 15
CC	FT	VARSPLIC 1 520
CC	FT	VARSPLIC 521 606
CC	FT	MISSED (IN ISOFORM 4).
CC	FT	DUEKINGQVHGVGGYDGEERKQNDKPNDSASPDHYH
CC	FT	DIQI -> MRRKKRFRVTVTIGEEGRNVOILMAIKVKA
CC	FT	KLDDLYDFG (IN ISOFORM 4).
CC	FT	I -> V (IN REF. 1).
CC	FT	L -> LVLK (IN REF. 1).
CC	FT	I -> N (IN REF. 1).
CC	FT	V -> A (IN REF. 1).
CC	FT	CONFLECT 1648 1648
CC	FT	SEQUENCE 1813 AA; 207250 MW; 6490319PIADEF02 C6C64;
CC	SO	Matched 8; Conservative 5; Mismatches 5; Indels 2; Gaps 1;
CC	Qy	10 NMAYGOVOO-NKEEDATN 27
CC	Dy	511 NMKSIIOEDNECDKNMCH 530
CC	RESULT 7	
CC	ID	GLPK_SVNY3 STANDARD; PRT; 495 AA.
CC	AC	P74260; (Ref. 35, Created)
CC	DT	01-JUN-1997 (Ref. 35, Sequence update)
CC	DE	15-JUN-2002 (Ref. 43, Last accession update)
CC	DE	Glycerol kinase [EC 2.7.1.30] (AMP:glycerol 3-phosphotransferase)
CC	GN	GLPK OR SLR167Z.
CC	GN	Synochrysis sp. (strain PCC 6803);
CC	OC	Bacteria; Cyanobacteria; Chroococcales; Synchrocystis.
CC	NCBI_TaxId	1148;
CC	RX	SEQUENCE FROM N.A.
CC	RX	MEDLINE:9706120; PubMed:8905231;
CC	RX	Kaneko T., Sato S., Kozaki H., Tanaka A., Asamizu E., Nakamura Y,
CC	RX	Miyajima N., Hirosewa M., Sugitara M., Sasamoto S., Kimura T.,
CC	RX	Hosouchi T., Matsuno A., Takouchi C., Wada T., Matsumae A.,
CC	RX	Ohtsuru S., Shimpo S., Takeuchi C., Wada T., Matsumae A.,
CC	RX	Yamada M., Tanabe M., Yakaba S. et al. of the unicellular cyanobacterium
CC	RX	Synechocystis sp. strain PCC6803 II genome sequence determination of the
CC	RX	entire genome and assignment of potential protein-coding regions.";
CC	RL	DNA Ref. 3:109-136(1996).
CC	CC	- FUNCTION: KEY ENZYME IN THE REGULATION OF GLYCEROL UPTAKE AND
CC	CC	METABOLISM.
CC	CC	- CATALYTIC ACTIVITY: ATP + glycerol = ADP + glycerol 3-phosphate.
CC	CC	- PATHWAY: Glycerol utilization, rate-limiting step.
CC	CC	- SIMILARITY: BELONGS TO THE ECOKINASE / GLUCOKINASE /
CC	CC	GLYCEROKINASE / ALDOSE KINASE FAMILY.
CC	CC	-----
CC	CC	This SWISS-Prot entry is copyright. It is produced through a collaboration-
CC	CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	CC	the European Bioinformatics Institute. There are no restrictions on its
CC	CC	use by non-profit institutions as long as its content is in no way

Query Match	Best local	Similarity	Score	DB 1	Length	1077
Matches	9	Conservative	7	Mismatches	10	Indels
DB	516	115MAVELEJENKWEKVEIRVDVIVS	541			
RESULT 9	ALP4_SCHPO	STANDARD	PRT	794 AA		
AD	ALP4_SCHPO					
AE	ALP4_SCHPO					
AF	ALP4_SCHPO					
AG	ALP4_SCHPO					
AH	ALP4_SCHPO					
AI	ALP4_SCHPO					
AL	ALP4_SCHPO					
AM	ALP4_SCHPO					
AN	ALP4_SCHPO					
AO	ALP4_SCHPO					
AP	ALP4_SCHPO					
AR	ALP4_SCHPO					
AS	ALP4_SCHPO					
AT	ALP4_SCHPO					
AV	ALP4_SCHPO					
AW	ALP4_SCHPO					
AX	ALP4_SCHPO					
AY	ALP4_SCHPO					
AZ	ALP4_SCHPO					
BA	ALP4_SCHPO					
BB	ALP4_SCHPO					
BC	ALP4_SCHPO					
BD	ALP4_SCHPO					
BE	ALP4_SCHPO					
BF	ALP4_SCHPO					
BG	ALP4_SCHPO					
BH	ALP4_SCHPO					
BI	ALP4_SCHPO					
BJ	ALP4_SCHPO					
BK	ALP4_SCHPO					
BL	ALP4_SCHPO					
BM	ALP4_SCHPO					
BN	ALP4_SCHPO					
BO	ALP4_SCHPO					
BP	ALP4_SCHPO					
BQ	ALP4_SCHPO					
BR	ALP4_SCHPO					
BS	ALP4_SCHPO					
BT	ALP4_SCHPO					
BU	ALP4_SCHPO					
BV	ALP4_SCHPO					
BW	ALP4_SCHPO					
BX	ALP4_SCHPO					
BY	ALP4_SCHPO					
BZ	ALP4_SCHPO					
CA	ALP4_SCHPO					
CB	ALP4_SCHPO					
CC	ALP4_SCHPO					
CD	ALP4_SCHPO					
CE	ALP4_SCHPO					
CF	ALP4_SCHPO					
CG	ALP4_SCHPO					
CH	ALP4_SCHPO					
CI	ALP4_SCHPO					
CJ	ALP4_SCHPO					
CK	ALP4_SCHPO					
CL	ALP4_SCHPO					
CM	ALP4_SCHPO					
CN	ALP4_SCHPO					
CO	ALP4_SCHPO					
CP	ALP4_SCHPO					
CQ	ALP4_SCHPO					
CR	ALP4_SCHPO					
CS	ALP4_SCHPO					
CT	ALP4_SCHPO					
CU	ALP4_SCHPO					
CV	ALP4_SCHPO					
CW	ALP4_SCHPO					
CX	ALP4_SCHPO					
CY	ALP4_SCHPO					
CA	ALP4_SCHPO					
CB	ALP4_SCHPO					
CC	ALP4_SCHPO					
CD	ALP4_SCHPO					
CE	ALP4_SCHPO					
CF	ALP4_SCHPO					
CG	ALP4_SCHPO					
CH	ALP4_SCHPO					
CI	ALP4_SCHPO					
CJ	ALP4_SCHPO					
CK	ALP4_SCHPO					
CL	ALP4_SCHPO					
CM	ALP4_SCHPO					
CN	ALP4_SCHPO					
CO	ALP4_SCHPO					
CP	ALP4_SCHPO					
CQ	ALP4_SCHPO					
CR	ALP4_SCHPO					
CS	ALP4_SCHPO					
CT	ALP4_SCHPO					

RC STRAIN-972; MEDLINE=21849401; PubMed=11895360;
 RA Med V., O'Neill S., Pearson D., Quail M.A., Rabbittowitch E.,
 RA Spoorer J., Peet N., Hayes U., Baker S., Baham C., Bowman S.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentes S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jagals K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Murphy K., Mullan L., Niblett D., Odeh C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbittowitch E.,
 RA Stelzrieder S., Smith M., Saunders D., Seeger S., Sharp S.,
 RA Taylor K., Taylor R.G., Tivy A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckreels G., Aert R., Robben J., Gymnopoulos B.,
 RA Meljens I., Vanterreels E., Rieger M., Schaefer M., Meulder-Auer S.,
 RA Gabel C., Fuchs M., Fyfe C., Leirner H., Reinhardt R., Pohl T.M.,
 RA Berger K., Zimmermann W., Medler H., Wendt R., Purnelle B.,
 RA Geller A., Gales S., Xeno S., Stock S., Molave V., Molave S.,
 RA Gellert M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
 RA Lucas M., Cruzado L., Jimenez J., Sanchez M., del Rey F., Bello J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,
 RA Cernelli L., Lowe T., McComble M.R., Paulsen I., Petashkin J.,
 RA Shipkoveki G.V., Usery D., Barrett B.G., Nurse P.;
 RT The genome sequence of *Schistosoma mansoni* pombae.
 CC Nucleotide sequence of the gamma tubule complex that is required
 CC for the regulation of both interphase microtubules and mitotic
 CC bipolar spindles.
 CC -1- SUBCELLULAR LOCATION: SPINDLE POLE BODY AND THE MICROTUBULE
 CC ORGANIZING CENTER (MTOC).
 CC -1- SIMILARITY: BELONGS TO THE GCP FAMILY.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announce/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL: A078627; CAB44757.1; -
 CC EMBL: A026664; BA077259.1; -
 CC Microtubules; Mitosis.
 CC SEQUENCE 784 AA; 90157 MW; E539CE217FFDA282 CRC64;
 SQ
 Query Match 36.04; Score 50; DB 1; Length 784;
 Best Local Similarity 31.24; Freq. No. 38;
 Matches 10; Conservative 8; Mismatches 8; Indels 6; Gaps 2;
 Oy 4 YRKXLMNAYGVQOKEDMAIEH--DWMM 32
 Db 592 FRYFLL---LRHVMQDLSNWSVQSKMSAVRL 620
 RESULT 10
 STD PART
 ID STYL PART STANDARD; PRT: 1574 AA.
 AC 062910; 062911; 089092;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DB Synapticin 1 (EC 3.1.3.56) (Synaptic Inositol-1,4,5-trisphosphate 5-
 DB Phosphatase 1).
 OS *Rattus norvegicus* (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_TaxID=10116;
 RX 11
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).

RC STRAIN-Sprague-Dawley; TISSUE=Brain;
 RA MEDLINE=96149250; PubMed=85521192;
 RA McPherson P.S., Garcia E.P., Stepien V.J., David C., Zhang X.,
 RA Grab D., Soehn M.S., Buerkleind R., Nemco T., De Camilli P.;
 RA Nature 379:353-357(1996).
 RL (2)
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RA TISSUE=Brain;
 RA MEDLINE=98374013; PubMed=9710239;
 RA Weschler R., Finn P.M., Ridley E., Parker P.J.;
 RT Identification and characterization of a novel splice variant of
 RT PPS8. Lett. 432:5-8(1999).
 RP ALTERNATIVE SPLICING.
 RP TISSUE=Brain;
 RA MEDLINE=96394655; PubMed=8798761;
 RA Ramjany A.R., McPherson P.S.;
 RT Identification of a novel alternative splicing gene that encodes a
 RL J Biol. Chem. 271:24856-24861(1996).
 CC -1- FUNCTION: INOSITOL-5-PHOSPHATASE WHICH HAS A ROLE IN CALYTHIN-
 CC MEDIATED ENDOCYTOSIS.
 CC -1- CATALYTIC ACTIVITY: D-xylo-inositol 1,4,5-trisphosphate + H(2)O =
 CC D-xylo-inositol 1,4-bisphosphate + phosphate.
 CC -1- SUBCELLULAR LOCATION: LOCALIZED MAINLY IN THE SOLUBLE FRACTION (BY
 CC HERB). ISOFORM 2/145 KDA AND ISOFORM 3/DELTA-SAC ARE PRODUCED BY
 CC ALTERNATIVE SPLICING. A STOP CODON IN POSITION 1309 IS SUPPRESSED
 CC IN ISOFORM 1 DUE TO ALTERNATIVE SPLICING. A 16 AA INSERT IS
 CC OPTIONALLY PRESENT IN THE THREE MAIN ISOFORMS.
 CC -1- TISSUE SPECIFICITY: ISOFORM 1 IS FOUND IN NEONATAL BRAIN AND IN A
 CC PREPOMNARY IN THE NERVOUS TISSUE. ISOFORM 2 IS EXPRESSED
 CC IN TISSUES AT WHICH LOWER LEVELS. ISOFORMS 1 AND 2 ARE DETECTED IN THE
 CC LUNG AND HEART. ISOFORM 1 IS EXPRESSED AT HIGHER LEVELS THAN
 CC ISOFORM 2 IN THE TESTIS AND LIVER AND BOTH ISOFORMS ARE NOT
 CC DETECTED IN THE SKELETAL MUSCLE. ISOFORM 3 WITH THE 16-AMINO-ACID
 CC INSERT IS ONLY FOUND IN THE BRAIN WHILE ISOFORM 3 WITHOUT THE 16-
 CC AMINO-ACID INSERT IS FOUND IN THE LUNG.
 CC -1- AMINO-ACID INSERTS FOUND IN THE LUNG.
 CC -1- SEREN WHILE AT E16 AND E18 ISOFORMS 1 AND 2 ARE SEREN IN THE ADULT
 CC BRAIN EXPRESSION OF ISOFORM 2 INCREASES DRASTICALLY AS COMPARED
 CC WITH ITS EXPRESSION IN EMBRYONIC BRAIN WHERE AS ISOFORM 1
 CC DECREASES TO UNDETECTABLE LEVELS.
 CC -1- DOMAIN: BINDS TO EPS15 (A CLATHRIN CONT-ASSOCIATED PROTEIN) VIA A
 CC C-TERMINAL C-TERMINAL PROLINE-RICH REGION MEDIATES BINDING TO A
 CC VARIETY OF SH3 DOMAIN-CONTAINING PROTEINS INCLUDING AMPHIPHILIN,
 CC SH34, SH39, SHP3 AND GRB2.
 CC -1- DOMAIN: SPLICING OF THE SAC1 DOMAIN DOES NOT ALTER THE CATALYTIC
 CC ACTIVITY OF SYNAPTOSIN 1.
 CC -1- SIMILARITY: IN THE CENTRAL SECTION; BELONGS TO THE INOSITOL-1,4,5-
 CC TRISPHOSPHATE 5-PHOSPHATASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SAC1 RECOGNITION MOTIF (RBM).
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announce/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL: U45479; AAB60525.1; -
 CC EMBL: U45479; AAB60526.1; -
 CC EMBL: A0006855; CA007267.1; ALT TERM.
 DR InterPro: IPR005135; Exo_Dphos.
 DR InterPro: IPR003100; IPFC.

DR InterPro: IPR000504; RNA_rec_mot.
 DR InterPro: IPR002013; Syla_N.
 DR Pfam: PF03383; Syla_N; 1.
 DR SMART: SM00232; Syla_N; 1.
 DR SMART: SM00126; IPR000000; 1.
 DR PROSITE: PS50275; SAC; 1.
 DR HydroLase; Alternative splicing; Repeat; Endocytosis; RNA-binding;
 KW Multigene family.
 FT DOMAIN 119 442 SAC.
 FT DOMAIN 590 937 RNA-BINDING (RGN).
 FT DOMAIN 900 1056 PRO-RICH.
 FT DOMAIN 1033 1036 POLY-SER.
 FT DOMAIN 1105 1110 POLY-PRO.
 FT DOMAIN 1123 1126 POLY-PRO.
 FT DOMAIN 1537 1545 POLY-PRO.
 FT DOMAIN 1401 1403 3 X 3 AA REPEATS OF N-P-F.
 FT REPEAT 1403 1403 1.
 FT REPEAT 1403 1403 1.
 FT REPEAT 1421 1423 3.
 FT VARSPLIC 1 400 MISSING (IN ISOFORM 3).
 FT VARSPLIC 1140 1155 MISSING (IN 16AA INSERTLESS ISOFORM).
 FT VARSPLIC 1309 1574 MISSING (IN ISOFORM 2).
 FT CONFLICT 588 588 G -> D (IN REF. 1; AAB60525).
 FT CONFLICT 1140 1140 G -> CG (IN REF. 2).
 FT CONFLICT 1155 1155 MISSING (IN REF. 1).
 FT SEQUENCE 1574 AA; 172860 MM; 3DBS:CB76B8ABD; CRC64;
 Query Match 26.0%; Score 50; DB 1; Length 1574;
 Best Local Similarity 27.6%; Pred. No. 83;
 Matches 8; Conservative 8; Mismatches 13; Indels 0; Gaps 0;
 Qy 3 SYRYKLNAVYQVQCNEDAMIEHDVVR 31
 Db 763 TIKDSDSDYDSCKTPHVDVLM 811

RESULT 11
 AC77 HUMAN STANDARD; PRT; 5430 AA.
 AC Q9UN2A; OSUKE0; GNUL29; Q9H540; Q70503;
 DT 1-OCT-2001 (Rel. 40) Created.
 DT 1-OCT-2001 (Rel. 40) Last annotation update.
 DE Actin cross-linking family protein 7 (Microphin) (Trabeculin-alpha)
 DE (620 kDa actin-binding protein) (AAB620).
 GN AC77 OR AAB620 OR KIA0465 OR KIA1251.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Canilata; Vertebrata; Euteleostomi;
 OC Mammalia; Scutellata; Primates; Catarrhini; Homiidae; Homo.
 ON NCBI; TaxId:9606;
 RN MEDLINE:20001959; PubMed:10529403;
 RA Okada T., Matsuda S., Nakatagawa S., Ichigotani Y., Iwahashi N.,
 RA Takahashi M., Ishigaki T., Hamaoguchi M.;
 RT Molecular cloning of microphin, a human homologue of Drosophila
 RT blood *Wt1b* related structural molecule family actin and dyactophila
 RT protein. Biophys. Res. Commun. 264:568-574(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE:20026884; PubMed:10559237;
 RA Sun Y., Zhang Y., Kretz S.-K., Ancalt D., Chang M.-S., Liu Y.,
 RA Sutherland R., Selig R., Griffin J.D., Feland L.H., Chen L.B.;
 RT Molecular cloning and characterization of human trabeculin-alpha, a
 RT protein that binds to actin. J. Biol. Chem. 274:33522-33530(1999).
 RU [3]
 RP SEQUENCE OF 868-2350 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE:20039619; PubMed:10574622;
 RA Nagase T., Ishikawa K.-I., Kikuno R., Hirose M., Nomura N.,

RA Ohara O.;
 RT Prediction of the coding sequences of unidentified human genes. XV.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for various proteins. J. Mol. Biol. 248:387-406(1997).
 RL NIA Res. 6137-345(1999).
 RN [4]
 RP SEQUENCE OF 1544-5057 FROM N.A.
 RA Cordy N.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 3734-5430 FROM N.A.
 RX MEDLINE:98116662; PubMed:9455484;
 RA Seki N., Ohira M., Nagase T., Ishikawa K.-I., Miyajima N.,
 RA Nakajima D., Nomura N., Ohara O.;
 RT "Characterization of cDNA clones in size-fractionated cDNA libraries
 RT from human brain." J.
 RL NIA Res. 4:345-349(1997).
 CC - FUNCTION: ACTIN-BINDING PROTEIN WHICH MAY PLAY A ROLE IN CROSS-
 CC - BINDING WITH OTHER CYTOSKELETAL PROTEINS. ALSO BINDS TO
 CC MICROTUBULES (BY SIMILARITY).
 CC - SUBCELLULAR LOCATION: Cytoplasmic.
 CC - TISSUE SPECIFICITY: UBQUITOUSLY EXPRESSED.
 CC - SIMILARITY: CONTAINS 1 ACTIN-BINDING DOMAIN.
 CC - SIMILARITY: CONTAINS 2 CALPOTIN-HOMOLOGY (CH) DOMAINS.
 CC - SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
 CC - SIMILARITY: CONTAINS 3 SH3 DOMAINS.
 CC - SIMILARITY: CONTAINS 3 REPEATS.
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.1so-stb.ch/announce/>
 CC or <http://www.ebi.ac.uk/seqdb/doc/1so-stb.html>).
 CC -----
 CC EMBL, AB029290; BAA8321.1; -;
 CC EMBL, AF141968; AAF06360.1; -;
 CC EMBL, AB033077; BAA8565.1; -;
 CC EMBL, AL137853; CAC15920.1; -;
 CC EMBL, AB079934; BAC13310.1; -;
 CC BSSB, Q02082.1; BAC13310.1; -;
 CC BSSB, Q02082.1; BAC13310.1; -;
 CC InterPro: IPR001569; Calpain-likn.
 CC InterPro: IPR001715; Calpain-likn.
 CC InterPro: IPR002048; EF-hand.
 CC InterPro: IPR003108; GAs2.
 CC InterPro: IPR001452; SH3.
 CC InterPro: IPR002817; Spectrin.
 CC Pfam: PF00335; Chhand; 2.
 CC Pfam: PF00435; spectrin; 36.
 CC Pfam: PF02187; GAs2; 1.
 CC ProDom: PD006012; EF-hand; 1.
 CC SMART: SM00033; EFh; 2.
 CC SMART: SM00054; EFh; 2.
 CC SMART: SM00233; GAs2; 1.
 CC PROSITE: PS00019; ACTININ_1.
 CC PROSITE: PS00020; ACTININ_2; FALSE_NEG.
 CC PROSITE: PS50021; CH; 2.
 CC PROSITE: PS50018; EF_HAND; 2.
 CC ACTIN-BINDING; Cytoskeleton; Calcium-binding; Repeat; SH3 domain.
 KW ACTIN-BINDING; Cytoskeleton; Calcium-binding; Repeat; SH3 domain.
 FT DOMAIN 78 125
 FT DOMAIN 194 295
 FT REPEAT 314 355 SPECTRIN 1.
 FT REPEAT 591 623 SPECTRIN 2.
 FT REPEAT 680 784 SPECTRIN 3.
 FT REPEAT 786 800 SPECTRIN 4.
 FT DOMAIN 871 923 SH3.

[illegible]

```

GN PX02-04.
GN Bacillus anthracis.
CC Plasmid pX02.
CC Bacillus anthracis.
CC Bacillus anthracis.
CC NCBI_TaxId=1392.
RP 1.
RP SEQUENCE FROM N.A.
RA Okimaka R.T., Cloud K., Hampton J., Hill K.K., Keim P., Lemke G.
RA Kramo S., Wanner D., Martinov Y., Svensson R., Tatum L.R.,
RA Brown A.E., Jackson P.O.
RA Submitted (09-1999) to the EMBL/Genbank/DBJ databases
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.1ab-sib.ch/announce/
CC or send an email to license@1ab-sib.ch.)
DR EMBL: AF186935.1;
DR Hypothetical protein, Plasmid.
KW SEQUENCE 95 AA; 11613 MW; 2C622DD9CFA66P CRC64;
SQ
Query Match 25.8%; Score 49.5; DB 1; Length 95;
Best Local Similarity 35.5%; Tm: No. 4.2;
Matches 11; Conservative 7; Mismatch 0; Indels 5; Gaps 1.
QY 1 PRSYRYKLVNMY-----QOVQNKKEDANIE 26
Db 29 RRYVMEIRYIMWFEQHFQYDSRNKEDKFE 59
-----
RESULT 13
DSDPBUCAI ID DSDPBUCAI STANDARD; PRT; 274 AA.
AC PE7232;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 2,3,4,5-tetrahydrocypidine-2-carboxylate N-succinyltransferase
DE EC 2.3.1.117 (tetrahydrocypidine N-succinyltransferase)
DE (EC:2.3.1.117) (tetrahydrocypidine N-succinyltransferase)
DE DAPD OR BU223 (tetrahydrocypidine N-succinyltransferase)
GN Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS Buchnera aphidicola)
OS symbiotic bacterium.
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
NCBI_TaxId=1180993;
CC EMBL: AF001118;
CC 11
CC SEQUENCE FROM N.A.
CC SHIMENOBU S., WATANABE S.
CC MEDLINE:20045113. PubMed:10993077.
CC SHIMENOBU S., Watanabe S., Hattori M., Sakaki Y., Ishikawa H.;
CC "Genome sequence of the endocellular bacterial symbiont of aphid
CC Buchnera sp. AFS."
CC Nature 407:81-86(2000).
RL -1- CATALYTIC ACTIVITY: Succinyl-CoA + 2,3,4,5-tetrahydrocypidine-2-
CC carboxylate + H2O = Succinyl-CoA + 2,3,4,5-tetrahydrocypidine-2-
CC carboxylate + H2O
CC -1- LYSINE FROM ASPARTATE SEMIALDEHYDE
CC -1- LYSINE FROM ASPARTATE SEMIALDEHYDE
CC -1- SIMILARITY: BELONGS TO THE CYSE/PLAMC (by similarity)
CC -1- SIMILARITY: BELONGS TO THE CYSE/PLAMC/PPA/NDOL FAMILY OF
CC ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LVI]-Q-X(4).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.1ab-sib.ch/announce/
CC or send an email to license@1ab-sib.ch.)
CC
CC EMBL: AP001118; BAB1944.1; -.

```

DR HSBP; P56220; 31DT.
DR InterPro; IPR001451; Hexapep. transf.
DR InterPro; IPR002325; Hexapep. 1.
DR TIGRFBP; TIGR00965; 1.
DR PROSITE; PS00101; HEXAPEP TRANSFERASES; 1.
KW Transferrase; Acyltransferase; Repeat; Lysine biosynthesis;
KW Diaminopimelate biosynthesis; Complete proteome.
SQ SEQUENCE 274 AA; 31184 MW; E559233E598A935 CRC64;
Query Match 25.84; Score 49.5; DB 1; Length 274;
Best Local Similarity 32.58; Pctd 40.10;
Matches 9; Conservative 6; Mismatches 8; Indels 1; Gaps 1;
7 KLNNAVQVQVQNEKEDAIHSHDW 30
||| : : : : :
DB KLNNGIIRISEKEDWTITHE-W 58
||| : : : : :
RESULT 14
ID XC73 METUA STANDARD; PRT; 350 AA.
AC Q56659;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein K01273.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxId=2190;
RN 1)
RP SEQUENCE FROM N.A.
RC STRAIN=AL-1; DSM 2463 / ATCC 33067;
RA Kohnen, S632399; PubMed=81687;
RA Bull C.G., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Overbeek A.R., Dougherty B.A., Tomb J.F., Adams R.A., Belch C.I.,
RA Rasmussen S.C., Karpman R., Loms D.J., Olson M.,
RA Scott J.L., Geoghegan N.S.M., Weinstock K.G., Merrick J.M., Glodek A.,
RA Ueberlack T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cordon M.D., Roberts K.M., Hutter R.A., Kaine B.P., Borodovsky M.,
RA White O., Olsen G.J., Smith T.A., Peterson J.D., Weinstock K.G.;
RT Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.
RT Science 273:1058-1073 (1996).
CC -1- SIMILARITY: TO M.JANNAVQVQVQNEKEDAIHSHDW (AC P21515).
CC 38.4 KDA PROTEIN IN SC55 5 REGION (AC P21515).
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF01595; AAB92929.1; -
CC F01595; AAB92929.1; -
CC DR InterPro; IPR002723; DUF43.
CC Pfam; PF01861; DUF43.1.
CC ProDom; PD011688; DUF43.1.
CC DR Hypothetical protein; Complete proteome.
KW SEQUENCE 350 AA; 40420 MW; 2F2018BF976C22B CRC64;
Query Match 25.84; Score 49.5; DB 1; Length 350;
Best Local Similarity 32.58; Pctd 40.10;
Matches 9; Conservative 6; Mismatches 7; Indels 9; Gaps 1;
9 LNNAVQVQVQNEKEDAIHSHDW 33
||| : : : : :
DB VNNQYIEETRAKLAFLPKKKEDIDWYASVRLIS 323
||| : : : : :
Query Match 25.84; Score 49.5; DB 1; Length 350;
Best Local Similarity 32.58; Pctd 40.10;
Matches 9; Conservative 6; Mismatches 7; Indels 9; Gaps 1;

RESULT 15
ID CYN CHEN STANDARD; PRT; 940 AA.
AC Q92387; Q9J045;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE VAL1-cRNA synthetase (EC 6.1.1.9) (Valine-cRNA ligase) (VALRS).
GN VALS OR CP00094 OR CP0060.
OS Bacteria; Chlamydiales; Chlamydia; Chlamydia pneumoniae.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxId=83558;
RN 1)
RP SEQUENCE FROM N.A.
RC STRAIN=CML029;
RA MEDLINE=9920606; PubMed=10192388;
RA Kaiman S., Mitchell M., Marathe R., Lammé C., Fan J., Hyman R.W.,
RA Kaiman S., Mitchell M., Marathe R., Lammé C., Fan J., Hyman R.W.,
RA Linher K., Weidman J., Knorr H., Craven B., Bowman C., Dodson R.,
RA Gullin M., Nelson M., Deboy R., Kolony J., McClary G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mogen and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406 (2000).
RN 1)
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Bunham R.C., Shen C., Gull J.S.R., Heideberg J.F., Tang S.,
RA Linher K., Weidman J., Knorr H., Craven B., Bowman C., Dodson R.,
RA Gullin M., Nelson M., Deboy R., Kolony J., McClary G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mogen and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406 (2000).
RN 1)
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20303049; PubMed=10871362;
RA Shiba T., Ishii K., Hattori M., Kihara S., Nakazawa T.,
RA Shiba T., Ishii K., Hattori M., Kihara S., Nakazawa T.,
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CML029 from USA.";
RL Nucleic Acids Res. 31:1-13 (2003).
CC -1- * L-valyl-cRNA (Val) AMP + L-valine + cRNA (Val) = AMP + diphosphate
CC
CC -1- SUBUNIT: MONOMER (by similarity).
CC
CC -1- SIMILARITY: BELONGS TO CLASS-1 AMINOACYL-TRNA SYNTHETASE FAMILY.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB01595; AAC1247.1; -
CC EMBL; AB01226; AAC3840.1; -
CC DR InterPro; IPR001421; Lysine synthetase.
DR HSBP; P06142; Lysine synthetase.
DR TIGR; CP06680; -
DR InterPro; IPR002300; cRNA-synth. 1a.
DR InterPro; IPR001412; cRNA-synth. 1.
DR InterPro; IPR002303; cRNA-synth. Val.
DR Pfam; PF00133; cRNA-synth. 1; 1.
DR Pfam; PF00133; cRNA-synth. 1; 1.
DR TIGRFBP; TIGR00422; VALRS.
KW PROSITE; PS00178; AA TRNA LIGASE I; 1.
KW Anticodon-cRNA synthetase; Protein biosynthesis; Ligase; AMP-binding;
KW Complete proteome.
FT SITE 47 "H1G" REGION.
FT SITE 564 "KNSG" REGION.
FT BINDING 567 ATP (BY SIMILARITY).

SQ SEQUENCE 940 AA: 107111 MM: 28054683F89D0404 CRC64;
 Query Match 25.8%; Score 48.5; DB 1; Length 940;
 Local Similarity 31.6%; Pct No. 54;
 Matches 12; Conservative 8; Mismatches 9; Indels 9; Gaps 2;
 QY 4 YR-YKLLN-----AVQGVQNKEDAMIEHDVWM 32
 DB 172 YRGYLVNMDPVLQTLADNDEVEHEKDGMLYIRYRM 209
 Search completed: March 26, 2003, 16:45:07
 Job time : 11.8837 secs

Db 1355 EVANENDNDNATHEHDVMA1SL 1377

RESULT 3

ES69631 galactosamine-containing minor telchotic acid biosynthesis 99ab - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #ext_change 15-Oct-1999

C:Accession: ES69631

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: ES69631

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: ES69631

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: ES69631

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: ES69631

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: ES69631

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: ES69631

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: ES69631

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: ES69631

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: ES69631

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: ES69631

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: ES69631

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: ES69631

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: ES69631

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: ES69631

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: ES69631

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: ES69631

hypothetical protein 2 - Microcystis aeruginosa

C:Species: Microcystis aeruginosa

C:Date: 01-Feb-1995 #sequence_revision 12-May-1995 #ext_change 22-Mar-2001

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

C:Accession: 1377

R:Name: F.; Ogawa, K.; Ogiwara, A.; Hosono, S.; Galizzi, A.; Alloni, G.; Azevedo, V.; Bertet

A:Gene: nar2369

Query Match

Best Local Similarity

Matches

DB

QY

RESULT 12

27222

hypothetical protein Y57G1C.10 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_rev10n 15-Oct-1999 #text_change 26-May-2000

C:Accession: T27222

R:NCBI:Tray: A

Submitted to the EMBL Data Library, September 1997

A:Accession: Z20330

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-444 <11>

A:Cross-References: EMBL:Z59281; PTD:CA816511.1; GSPDB:GN00022; CESP:Y57G1C.10

A:Experimental source: clone Y57G1C.10

C:Genetic: CESP:Y57G1C.10

A:Map position: 4

A:Interon: 332/1; 416/3

C:Superfamily: human GDP dissociation inhibitor XAP-4

Query Match

Best Local Similarity

Matches

DB

QY

RESULT 13

57595

probable L-xylulokinase (EC 2.7.1.53) - *Synechocystis* sp. (strain PCC 6803)C:Species: *Synechocystis* sp.

A:Date: 10-Sep-1999 #sequence_rev10n 10-Sep-1999 #text_change 16-Jun-2000

C:Accession: S75955

R:Nameo, T.; Sato, S.; Kocant, H.; Tanaka, A.; Asami, R.; Nakamura, Y.; Miyajima, N.;

DNA Res. 8, 109-116, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

A:Reference number: S74322; MUID:97061201; PMID:8905231

C33260

hypothetical protein PA079 [imported] - *Pseudomonas aeruginosa* (strain PA01)C:Species: *Pseudomonas aeruginosa*

C:Date: 14-Dec-2000 #sequence_rev10n 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: C33260

R:Scriver, C.K.; Phan, X.Q.; Ewlin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; By

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardis, J.; Lim,

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho

A:Accession: C33260

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-793 <S70>

A:Cross-References: GB:AE004732; GB:AE004091; NID:994949186; PTD:AA00467.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetic: PA079

A:Gene: PA079

Query Match

Best Local Similarity

Matches

DB

QY

RESULT 15

AG2434

hypothetical protein al1503 [imported] - *Neisseria* sp. (strain PCC 7120)C:Species: *Neisseria* sp.

C:Date: 14-Dec-2000 #sequence_rev10n 14-Dec-2000 #text_change 30-Jun-2002

C:Accession: AG2434

R:Nameo, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Iriyuchi,

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the filamentous Nitrogen-fixing Cyanobacterium *Ana*

A:Reference number: AB1807; MUID:21595285; PMID:11759640

A:Cross-References: GB:BA000019; PTD:BA076730.1; PTD:G1734169; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetic: A1503

A:Superfamily: *Synechocystis* hypothetical protein al1203

Query Match

Best Local Similarity

Matches

DB

QY

RESULT 16

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

11

Search completed: March 26, 2003, 16:50:40
 Job time : 21.3721 secs

Genome version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 16:49:57 / Search time 22.9102 Seconds
(Without alignments) 87.073 Million cell updates/sec

Title: US-10-010-667a-20

Sequence: 1 RSRYRYKLMAVYQVQONKEDAKIETHDWMKEI 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 237916 seqs, 58723574 residues

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA.*

1: /cgn2_6/prodatea/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodatea/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodatea/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodatea/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodatea/1/pubpaa/US07_PUBCOMB.pep.*
6: /cgn2_6/prodatea/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodatea/1/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/prodatea/1/pubpaa/US09_NEW_PUB.pep.*
9: /cgn2_6/prodatea/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/prodatea/1/pubpaa/US10_NEW_PUB.pep.*
11: /cgn2_6/prodatea/1/pubpaa/US10_PUBCOMB.pep.*
12: /cgn2_6/prodatea/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/prodatea/1/pubpaa/US06_PUBCOMB.pep.*
14: /cgn2_6/prodatea/1/pubpaa/US06_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score distribution, and is derived by analysis of the score score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	192	100.0	34	US-10-011-095-20	Sequence 20, App1
2	192	100.0	34	US-10-010-667a-20	Sequence 20, App1
3	192	100.0	267	US-09-147-835a-50	Sequence 50, App1
4	192	100.0	267	US-09-147-835a-51	Sequence 51, App1
5	192	100.0	339	US-09-802-520-179	Sequence 11, App1
6	192	100.0	339	US-09-802-520-179	Sequence 11, App1
7	192	100.0	339	US-09-895-814-879	Sequence 879, App
8	192	100.0	339	US-09-895-814-879	Sequence 879, App
9	192	100.0	339	US-10-011-095-2	Sequence 2, App1
10	192	100.0	339	US-10-010-667a-2	Sequence 2, App1
11	192	100.0	339	US-08-159-131-879	Sequence 879, App
12	192	100.0	339	US-08-159-131-879	Sequence 879, App
13	192	100.0	339	US-08-159-131-879	Sequence 879, App
14	192	100.0	339	US-08-159-131-879	Sequence 879, App
15	192	100.0	339	US-08-159-131-879	Sequence 879, App
16	192	100.0	339	US-08-159-131-879	Sequence 879, App
17	192	100.0	339	US-08-159-131-879	Sequence 879, App
18	192	100.0	339	US-08-159-131-879	Sequence 879, App
19	192	100.0	339	US-08-159-131-879	Sequence 879, App

20	85	44.3	15	US-10-011-095-32	Sequence 32, App1
21	85	44.3	15	US-10-010-667a-32	Sequence 32, App1
22	85	44.3	15	US-09-147-835a-50	Sequence 50, App1
23	85	44.3	15	US-09-147-835a-51	Sequence 51, App1
24	51.5	26.8	491	US-09-802-520-179	Sequence 11, App1
25	51	26.6	487	US-09-802-520-179	Sequence 11, App1
26	50	26.0	370	US-09-895-814-879	Sequence 879, App
27	50	26.0	370	US-09-895-814-879	Sequence 879, App
28	50	26.0	370	US-09-895-814-879	Sequence 879, App
29	48.5	25.3	151	US-10-010-667a-2	Sequence 2, App1
30	48.5	25.3	151	US-10-010-667a-2	Sequence 2, App1
31	48.5	25.3	151	US-10-010-667a-2	Sequence 2, App1
32	48.5	25.3	151	US-10-010-667a-2	Sequence 2, App1
33	48.5	25.3	151	US-10-010-667a-2	Sequence 2, App1
34	48.5	25.3	151	US-10-010-667a-2	Sequence 2, App1
35	47.5	24.7	4150	US-09-808-880-2	Sequence 2, App1
36	47	24.5	274	US-09-739-907-171	Sequence 171, App1
37	47	24.5	274	US-09-739-907-171	Sequence 171, App1
38	47	24.5	274	US-09-739-907-171	Sequence 171, App1
39	47	24.5	274	US-09-739-907-171	Sequence 171, App1
40	47	24.5	274	US-09-739-907-171	Sequence 171, App1
41	46.5	24.2	135	US-09-764-664-1191	Sequence 1191, App
42	46.5	24.2	135	US-09-764-664-1191	Sequence 1191, App
43	46.5	24.2	135	US-09-764-664-1191	Sequence 1191, App
44	46.5	24.2	135	US-09-764-664-1191	Sequence 1191, App
45	46.5	24.2	135	US-09-764-664-1191	Sequence 1191, App

ALIGNMENTS

RESULT 1
US-10-011-095-20
Sequence 20, Application US/10011095
US-10-010-667a-20
Sequence 20, Application US/1001062A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Hubert, Rene S.
APPLICANT: Leong, Kahen
APPLICANT: Raitano, Arthur B.
APPLICANT: Salzman, Douglas C.
APPLICANT: Stetson, Steven
FILE REFERENCE: 51152001610
CURRENT APPLICATION NUMBER: US/10/011.095
PRIOR FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: 09/323.873
PRIOR FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: 60/087.520
PRIOR FILING DATE: 1998-06-01
PRIOR APPLICATION NUMBER: 60/091.183
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 20
LENGTH: 34
TYPE: PRT
PROTEIN: Artificial Sequence
OTHER INFORMATION: STAP-1 peptide
US-10-011-095-20
Query Match 100.0%; Score 192; DB 9; Length 34;
Best Local Similarity 100.0%; Pred. No. 1,1e-19;
Matches 34; Mismatches 0; Indels 0; Gaps 0;
CY 1 RSRYRYKLMAVYQVQONKEDAKIETHDWMKEI 34
DB 1 RSRYRYKLMAVYQVQONKEDAKIETHDWMKEI 34
RESULT 2
US-10-010-667a-20

```

/ Sequence 20, Application US/10010667a
/ Publication No. US20030055217A1
/ GENERAL INFORMATION:
/ APPLICANT: Alar, Daniel S.
/ APPLICANT: Alar, Daniel S.
/ APPLICANT: Raitano, Arthur B.
/ APPLICANT: Saffran, Douglas C.
/ APPLICANT: Mitchell, Steve Chappell
/ TITLE OF INVENTION: NOVEL SERPENTIN TRANSMEMBRANE ANTIGENS
/ FILE REFERENCE: HVS-37CIP
/ CURRENT APPLICATION NUMBER: US/10/010,667a
/ PRIOR FILING DATE: 2001-12-06
/ PRIOR APPLICATION NUMBER: 09/323,873
/ PRIOR FILING DATE: 1999-06-01
/ PRIOR APPLICATION NUMBER: 60/087,520
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/091,183
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: ParesSeq for Windows Version 4.0
/ SEQ ID NO 20
/ LENGTH: 34
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURES:
/ INFORMATION: STAMP-1 peptide
US-10-010-667a-20

Query Match          100.0%; Score 192; DB 9; Length 34;
Best Local Similarity 100.0%; Pred. No. 1,1e-19;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RRSRYKILMAVYQOONKEDMIEHDVWMEI 34
Db 1 RRSRYKILMAVYQOONKEDMIEHDVWMEI 34

RESULT 3
US-09-747-835a-50
/ Sequence 50, Application US/09747835a
/ Publication No. US20030055217A1
/ GENERAL INFORMATION:
/ APPLICANT: Yamazaki, Victoria
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Zhou, Ping
/ APPLICANT: Wang, Duntui
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Peiyun
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Dimaecio, Rodolfo T
/ TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
/ FILE REFERENCE: HVS-37CIP
/ CURRENT APPLICATION NUMBER: US/09/747,835a
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: US 09/653,450
/ PRIOR FILING DATE: 2000-12-04
/ PRIOR APPLICATION NUMBER: US 09/729,739
/ PRIOR FILING DATE: 2000-08-31
/ PRIOR APPLICATION NUMBER: US 09/650,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: US 09/558,042
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: US 09/552,317
/ PRIOR FILING DATE: 2000-01-21
/ PRIOR APPLICATION NUMBER: US 09/488,725
/ NUMBER OF SEQ ID NOS: 63
/ SOFTWARE: Patentin version 3.0
/ SEQ ID NO 50

```

```

/ LENGTH: 267
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-747-835a-50

Query Match          100.0%; Score 192; DB 10; Length 267;
Best Local Similarity 100.0%; Pred. No. 1e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RRSRYKILMAVYQOONKEDMIEHDVWMEI 34
Db 119 RRSRYKILMAVYQOONKEDMIEHDVWMEI 152

RESULT 4
US-09-747-835a-51
/ Sequence 51, Application US/09747835a
/ Publication No. US20020146692A1
/ GENERAL INFORMATION:
/ APPLICANT: Yamazaki, Victoria
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Zhou, Ping
/ APPLICANT: Wang, Duntui
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Peiyun
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Dimaecio, Rodolfo T
/ TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
/ FILE REFERENCE: HVS-37CIP
/ CURRENT APPLICATION NUMBER: US/09/747,835a
/ PRIOR FILING DATE: 2002-03-08/729,739
/ PRIOR APPLICATION NUMBER: US 09/653,450
/ PRIOR FILING DATE: 2000-12-04
/ PRIOR APPLICATION NUMBER: US 09/653,450
/ PRIOR FILING DATE: 2000-08-31
/ PRIOR APPLICATION NUMBER: US 09/650,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: US 09/558,042
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: US 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: US 09/488,725
/ NUMBER OF SEQ ID NOS: 63
/ SOFTWARE: Patentin version 3.0
/ SEQ ID NO 51
/ LENGTH: 267
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-747-835a-51

Query Match          100.0%; Score 192; DB 10; Length 267;
Best Local Similarity 100.0%; Pred. No. 1e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RRSRYKILMAVYQOONKEDMIEHDVWMEI 34
Db 119 RRSRYKILMAVYQOONKEDMIEHDVWMEI 152

RESULT 5
US-10-012-896-879
/ Sequence 879, Application US/10012896
/ Publication No. US20030055217A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Dillon, David C.
/ APPLICANT: Mitchell, Jennifer L.
/ APPLICANT: Harlocker, Susan L.
/ APPLICANT: Jiang, Yuguu
/ APPLICANT: Kalos, Michael D.

```

```

/ APPLICANT: Retter, Marc W.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darrick
/ APPLICANT: Li, Samuel X.
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skelky, Yasir A.W.
/ APPLICANT: Hepler, William T.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Houghton, Raymond L.
/ APPLICANT: Vinals de Basols, Carlota
/ APPLICANT: Foy, Teresa
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Manande, Yoshinori
/ APPLICANT: Resigner, Modelaine
/ APPLICANT: Resigner, Modelaine
/ TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
/ FILE REFERENCE: 210121.427C27
/ CURRENT APPLICATION NUMBER: US/10/012, 896
/ CURRENT FILING DATE: 2001-12-10
/ NUMBER OF SEQ ID NOS: 1011
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 979
/ SEQ ID NO 980
/ TYPE: PAT
/ ORGANISM: Homo sapiens
US-10-012-896-879

```

```

Query Match          100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 1,3e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNMAVYQVQONKEDAMTEHDVWMEI 34
|||||
DB 185 RRSRYKLLNMAVYQVQONKEDAMTEHDVWMEI 218

```

```

RESULT 6
US-09-802-520-11
/ Sequence 520; Application US/09802520
/ Publication No. US20020187472A1
/ GENERAL INFORMATION:
/ APPLICANT: Lal, Preeti
/ APPLICANT: Paris, Mary
/ APPLICANT: Chen, Hui-Mei
/ APPLICANT: Ison, Craig H.
/ TITLE OF INVENTION: STEAR-RELATED PROTEIN
/ FILE REFERENCE: 210121.427C27
/ CURRENT APPLICATION NUMBER: US/09/802, 520
/ CURRENT FILING DATE: 2001-03-09
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: PERL Program
/ SEQ ID NO 11
/ LENGTH: 339
/ TYPE: PAT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. US20020187472A1 96572948
US-09-802-520-11

```

```

Query Match          100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 1,3e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNMAVYQVQONKEDAMTEHDVWMEI 34
|||||
DB 185 RRSRYKLLNMAVYQVQONKEDAMTEHDVWMEI 218

```

```

RESULT 7
US-09-895-793-879
/ Sequence 879; Application US/09895793
/ Publication No. US20020192763A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Dillon, Devin C.
/ APPLICANT: Mitchem, Jennifer L.
/ APPLICANT: Harlocke, Susan L.
/ APPLICANT: Wang, Michael D.
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darrick
/ APPLICANT: Li, Samuel X.
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skelky, Yasir A.W.
/ APPLICANT: Hepler, William T.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Hural, John
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Houghton, Raymond L.
/ APPLICANT: Vinals de Basols, Carlota
/ APPLICANT: Foy, Teresa
/ APPLICANT: Fanger, Gary R.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.534C2
/ CURRENT APPLICATION NUMBER: US/09/895, 793
/ CURRENT FILING DATE: 2001-06-29
/ NUMBER OF SEQ ID NOS: 982
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 879
/ SEQ ID NO 879
/ LENGTH: 339
/ TYPE: PAT
/ ORGANISM: Homo sapiens
US-09-895-793-879

```

```

Query Match          100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 1,3e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNMAVYQVQONKEDAMTEHDVWMEI 34
|||||
DB 185 RRSRYKLLNMAVYQVQONKEDAMTEHDVWMEI 218

```

```

RESULT 8
US-09-895-814-879
/ Sequence 879; Application US/09895814
/ Publication No. US2002019236A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Dillon, Devin C.
/ APPLICANT: Mitchem, Jennifer L.
/ APPLICANT: Harlocke, Susan L.
/ APPLICANT: Wang, Michael D.
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darrick
/ APPLICANT: Li, Samuel X.
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skelky, Yasir A.W.
/ APPLICANT: Hepler, William T.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Hural, John
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Houghton, Raymond L.

```

```

/ APPLICANT: Vinals de Baseols, Carlota
/ APPLICANT: Foy, Teresa
/ APPLICANT: Fanger, Gary R.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
/ FILE REFERENCE: 210121.427C26
/ CURRENT APPLICATION NUMBER: US/09/895, 814
/ CURRENT FILING DATE: 2001-06-29
/ NUMBER OF SEQ ID NOS: 990
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 979
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-895-814-879

Query Match      100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 1,3e-18;
Matches 34; Conservative 0; Mismatches 0; Gaps 0;

Cy 1 RRSRYKLLMAYQOONKEDAMIEHDVWMEI 34
|||||
Db 185 RRSRYKLLMAYQOONKEDAMIEHDVWMEI 218

RESULT 9
US-10-011-095-2
/ Sequence 2: Application US/10011095
/ Publication No. US20030045682A1
/ GENERAL INFORMATION:
/ APPLICANT: Aear, Daniel
/ APPLICANT: Hubert, Rene S.
/ APPLICANT: Leong, Arthur B.
/ APPLICANT: Leong, Kahen
/ APPLICANT: Mitchell, Steven Chappell
/ TITLE OF INVENTION: ANTIBODIES IMMUNOSPECIFIC FOR STAP1 (AS AMENDED)
/ FILE REFERENCE: 511582001610
/ CURRENT APPLICATION NUMBER: US/10/011, 095
/ CURRENT FILING DATE: 2001-12-06
/ PRIOR APPLICATION NUMBER: 09/323,873
/ PRIOR FILING DATE: 1999-06-01
/ PRIOR FILING DATE: 1998-06-01/087,520
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/091,183
/ PRIOR FILING DATE: 1998-06-30
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ OTHER INFORMATION: DNA
US-10-011-095-2

Query Match      100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 1,3e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RRSRYKLLMAYQOONKEDAMIEHDVWMEI 34
|||||
Db 185 RRSRYKLLMAYQOONKEDAMIEHDVWMEI 218

```

```

/ APPLICANT: Raitano, Arthur B.
/ APPLICANT: Saitan, Douglas C.
/ APPLICANT: Mitchell, Steven Chappell
/ TITLE OF INVENTION: NOVEL SEQUENCE TRANSMEMBRANE ANTIGENS
/ TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USBS THEREOF
/ FILE REFERENCE: 511582001601
/ CURRENT APPLICATION NUMBER: US/10/010,667A
/ CURRENT FILING DATE: 2001-12-06
/ PRIOR APPLICATION NUMBER: 09/323,873
/ PRIOR FILING DATE: 1999-06-01
/ PRIOR FILING DATE: 1998-06-01/087,520
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/091,183
/ PRIOR FILING DATE: 1998-06-30
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-010-667A-2

Query Match      100.0%; Score 192; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 1,3e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RRSRYKLLMAYQOONKEDAMIEHDVWMEI 34
|||||
Db 185 RRSRYKLLMAYQOONKEDAMIEHDVWMEI 218

RESULT 11
US-09-755-143-879
/ Sequence 879: Application US/09759143
/ Publication No. US200202248A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jianshun
/ APPLICANT: Dillon, David C.
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: Harlocker, Susan L.
/ APPLICANT: Jiang, Yugu
/ APPLICANT: Kellum, Robert A.
/ APPLICANT: Kellum, Michael D.
/ APPLICANT: Rafter, Gary R.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Derrick
/ APPLICANT: M. Samad
/ APPLICANT: Wang, William
/ APPLICANT: Skeiky, Yassir A.W.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
/ FILE REFERENCE: 210121.427C23
/ CURRENT APPLICATION NUMBER: US/09/759,143
/ CURRENT FILING DATE: 2002-01-12
/ NUMBER OF SEQ ID NOS: 934
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 879
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-755-143-879

Query Match      100.0%; Score 192; DB 10; Length 339;
Best Local Similarity 100.0%; Pred. No. 1,3e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RRSRYKLLMAYQOONKEDAMIEHDVWMEI 34
|||||
Db 185 RRSRYKLLMAYQOONKEDAMIEHDVWMEI 218

```

```

RESULT 12
US-09-780-659-879
Sequence 879, Application US/09780669
Patent No. US20020051977A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Devin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Haddock, Susan L.
APPLICANT: Haddock, Susan L.
APPLICANT: Henderson, Robert A.
APPLICANT: Katos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedrick, Thomas S.
APPLICANT: Lister, Patrick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aljun
APPLICANT: Skeiky, Yaeli A.W.
APPLICANT: Hepler, William
APPLICANT: Hurai, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
TITLE OF INVENTION: DIAGNOSIS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C24
CURRENT FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 943
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 939
SEQ ID NO 938
SEQ ID NO 937
TYPE: PRT
ORGANISM: Homo sapiens
US-09-780-659-879

Query Match 100.0%; Score 192; DB 10; Length 339;
Best Local Similarity 100.0%; Pred. No. 1,3e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRSRYKLLNMAVOOVQONKEDAMIEHWMEI 34
|||||
DB 185 RRSRYKLLNMAVOOVQONKEDAMIEHWMEI 218

RESULT 13
US-09-822-827-879
Sequence 879, Application US/09822827
Patent No. US20020081680A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.534C1
CURRENT APPLICATION NUMBER: US/09/822,827
FILE REFERENCE: 210121.427C24
NUMBER OF SEQ ID NOS: 982
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 879
LENGTH: 339
TYPE: PRT
ORGANISM: Homo sapiens
US-09-822-827-879

Query Match 100.0%; Score 192; DB 10; Length 339;
Best Local Similarity 100.0%; Pred. No. 1,3e-18;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRSRYKLLNMAVOOVQONKEDAMIEHWMEI 34
|||||
DB 185 RRSRYKLLNMAVOOVQONKEDAMIEHWMEI 218

```

```

DB 185 RRSRYKLLNMAVOOVQONKEDAMIEHWMEI 218
|||||

RESULT 14
US-09-963-896-1
Sequence 1, Application US/09963896
Patent No. US20020102585A1
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Lal, Preeti
APPLICANT: Lal, Preeti
TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSER: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION NUMBER: US/09/963,896
FILING DATE: 26-SEP-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/397,558
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION: C.
REGISTRATION NUMBER: 39,112
REFERENCE/DOCKET NUMBER: PR-0527 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 845-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 339
TYPE: amino acid
STRANDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSTUT10
CLONE: 1691243
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-963-896-1

Query Match 55.7%; Score 107; DB 10; Length 141;
Best Local Similarity 58.8%; Pred. No. 1,8e-07;
Matches 20; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 13 RRSRYKLLNMAVOOVQONKEDAMIEHWMEI 34
|||||
DB 13 RRSRYKLLNMAVOOVQONKEDAMIEHWMEI 46

RESULT 15
US-10-011-095-8
Sequence 8, Application US/10011095
Patent No. US20030045682A1
GENERAL INFORMATION:
APPLICANT: Rabinowitz, Daniel
APPLICANT: Hubert, Rene S.
APPLICANT: Leong, Kahan
APPLICANT: Raitano, Arthur B.
APPLICANT: Saffari, Douglas C.
APPLICANT: Mitchell, Steve Chappeil
TITLE OF INVENTION: ANTIBODIES IMMUNOSPECIFIC FOR STEAP1 (AS AMENDED)

```



```

/ TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
/ TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USGS THEREOF
/ FILE REFERENCE: 129.16US02
/ INVENTOR: US/09/323,873A
/ CURRENT FILING DATE: 1999-06-01
/ PRIOR APPLICATION NUMBER: 60/087,520
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/091,183
/ PRIOR FILING DATE: 1998-06-30
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: PatsSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-09-323-873A-2

Query Match 100.0%; Score 192; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 4, 9e-10;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLNMAVOOVQONKEDAMIEHDVWMEI 34
DB 185 RRSRYKLLNMAVOOVQONKEDAMIEHDVWMEI 218

RESULT 3
US-09-083-521-1
/ Sequence 1, Application US/09083521
/ Patent No. 6048970
/ GENERAL INFORMATION:
/ APPLICANT: Lal, Preeti
/ APPLICANT: Guegler, Karl J.
/ APPLICANT: Guegler, Neil
/ TITLE OF INVENTION: FASTSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS
/ NUMBER OF SEQUENCES: 7
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
/ STREET: 3174 PORTER DRIVE
/ CITY: PALO ALTO
/ STATE: CALIFORNIA
/ ZIP: 94304
/ COUNTRY: USA
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ FILING DATE: Herewith
/ CLASSIFICATION: Herewith
/ ATTORNEY/AGENT INFORMATION:
/ NAME: CERRONE, MICHAEL C.
/ REGISTRATION NUMBER: 39,132
/ REFERENCE/DOCKET NUMBER: PF-0527 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 945-4165
/ TELEFAX: (650) 945-4166
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 141 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ INSDCAT: 1
/ LIBRARY: PROSTUT10
/ CLONE: 1691243
US-09-083-521-1

Query Match 55.7%; Score 107; DB 3; Length 141;
Best Local Similarity 58.8%; Pred. No. 1e-07; 9; Indels 0; Gaps 0;
Matches 20; Conservative 5; Mismatches 0;

```

```

QY 1 RRSRYKLLNMAVOOVQONKEDAMIEHDVWMEI 34
DB 13 RRSRYKLLNMAVOOVQONKEDAMIEHDVWMEI 46

RESULT 4
US-09-323-873A-8
/ Sequence 8, Application US/0932873A
/ Patent No. 6329303
/ GENERAL INFORMATION:
/ APPLICANT: Daniel R. Afar
/ APPLICANT: Rene S. Hubert
/ APPLICANT: Kahan Leong
/ APPLICANT: Arthur B. Raitano
/ APPLICANT: Douglas C. Saffran
/ APPLICANT: Steve Chappell Mitchell
/ TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
/ TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USGS THEREOF
/ FILE REFERENCE: 129.16US02
/ CURRENT APPLICATION NUMBER: US/09/323,873A
/ CURRENT FILING DATE: 1999-06-01
/ PRIOR APPLICATION NUMBER: 60/087,520
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/091,183
/ PRIOR FILING DATE: 1998-06-30
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: PatsSeq for Windows Version 4.0
/ SEQ ID NO 8
/ LENGTH: 173
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-09-323-873A-8

Query Match 55.7%; Score 107; DB 4; Length 173;
Best Local Similarity 58.8%; Pred. No. 1, 3e-07;
Matches 20; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 RRSRYKLLNMAVOOVQONKEDAMIEHDVWMEI 34
DB 81 RRSRYKLLNMAVOOVQONKEDAMIEHDVWMEI 114

RESULT 5
US-09-323-873A-32
/ Sequence 32, Application US/0932873A
/ Patent No. 6329303
/ GENERAL INFORMATION:
/ APPLICANT: Daniel R. Afar
/ APPLICANT: Rene S. Hubert
/ APPLICANT: Kahan Leong
/ APPLICANT: Arthur B. Raitano
/ APPLICANT: Douglas C. Saffran
/ APPLICANT: Steve Chappell Mitchell
/ TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
/ TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USGS THEREOF
/ FILE REFERENCE: 129.16US02
/ CURRENT APPLICATION NUMBER: US/09/323,873A
/ CURRENT FILING DATE: 1999-06-01
/ PRIOR APPLICATION NUMBER: 60/087,520
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/091,183
/ PRIOR FILING DATE: 1998-06-30
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: PatsSeq for Windows Version 4.0
/ SEQ ID NO 32
/ LENGTH: 15
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-09-323-873A-32

Query Match 44.3%; Score 85; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 5, 1e-06;

```


Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 YOUNGKEMM187 27
|||||||
Db 1 YOUNGKEMM187 15

RESULT 6
US-09-149-476-436
Sequence 436, Application US/09149476
Patent No. 6420322
INVENTOR: YOUNG, KENNETH
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 human secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149, 476
EARLIER FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,521
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,526
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,533
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,517
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845

```

/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,892
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,892
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/047,595
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/047,599
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,588
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,585
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,586
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,590
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,594
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,589
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,593
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,614
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/043,576
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,578
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,501
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/043,670
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/056,632
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,664
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,876
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,881
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,909
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,875
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,862
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,887
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,908
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/048,964
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/057,650
/ EARLIER FILING DATE: 1997-09-05
/ EARLIER APPLICATION NUMBER: 60/056,884
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/057,669
/ EARLIER FILING DATE: 1997-09-05
/ EARLIER APPLICATION NUMBER: 60/049,610
/ EARLIER FILING DATE: 1997-06-13
/ EARLIER APPLICATION NUMBER: 60/061,060
/ EARLIER FILING DATE: 1997-10-02

```

```

Query Match 26.04; Score 50; DB 4; Length 370;
Local Similarity 20.17; Pctd. No. 0;
Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

```

```

QY 5 RYKLNAYVOOQ--VOOKEDAWIHDVWR 31
DB 191 RQRKNDALRLELKEFANPDVWR 217

```

```

RESULT 7
US-08-560-005-10
/ Sequence 10; Application US/08560005
/ Patent No. 6236848
/ GENERAL INFORMATION:
/ APPLICANT: Pot, David A.
/ APPLICANT: Williams, Lewis T.
/ APPLICANT: Jefferson, Anne Bennett
/ TITLE OF INVENTION: No. 60013561 Grp2 Associating Protein and Nucleic
/ NUMBER OF SEQUENCES: 10
/ NUMBER OF SEQUENCES: 10
/ ADDRESS: Townsend and Townsend and Crew
/ STREET: One Market Plaza, Steuart Tower, Suite 2000
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/560,005
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Dow, Karen B.
/ REGISTRATION NUMBER: 29,684
/ REFERENCE/DOCKET NUMBER: 2307K-0624000
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-762-9420
/ TELEFAX: 415-762-9422
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 654 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULAR TYPE: protein
/ EXPRESSION:
/ NAME/KEY: Region
/ LOCATION: 1..654
/ OTHER INFORMATION: /note="consensus"
US-08-560-005-10

```

```

Query Match 25.34; Score 48.5; DB 3; Length 654;
Best Local Similarity 25.04; Pctd. No. 0;
Matches 8; Conservative 12; Mismatches 9; Indels 3; Gaps 1;
QY 3 SYRYKLNAYVOO--VOOKEDAWIHDVWR 31
DB 348 TYXRGNDVAVYKQRDSEKTRVPMCDRLMK 379

```

```

RESULT 8
US-08-560-005-10
/ Sequence 10; Application US/09418540
/ Patent No. 6236848
/ GENERAL INFORMATION:
/ APPLICANT: Pot, David A.
/ APPLICANT: Williams, Lewis T.
/ APPLICANT: Jefferson, Anne Bennett
/ APPLICANT: Majerus, Philip M.
/ TITLE OF INVENTION: No. 60013561 Grp2 Associating Protein and Nucleic
/ NUMBER OF SEQUENCES: 10
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSES: Townsend and Townsend and Crew
/ STREET: One Market Plaza, Steuart Tower, Suite 2000
/ CITY: San Francisco
/ STATE: California

```

/ COUNTRY: USA
 / ZIP: 94105
 / COMPUTER: READABLE FORM: disk
 / MEDIA: Tapes, 1/2 inch
 / OPERATING SYSTEM: IBM PC Compatible
 / SOFTWARE: Patent In Release #1.0, Version #1.25
 / CURRENT APPLICATION DATA:
 / PENDING DATE: 14-OCT-1999
 / PUBLICATION DATA:
 / APPLICATION NUMBER: US 08/560,005
 / FILING DATE: 17-NOV-1995
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Dow, Karen B.
 / REGISTRATION NUMBER: 29,684
 / REFERENCE/DOCKET NUMBER: 2307K-0624000
 / TELEPHONE: 415-326-3400
 / TELEFAX: 415-326-2422
 / INFORMATION FOR SEQ ID NO: 10:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 654 amino acids
 / TYPE: amino acid
 / TOPOLOGY: linear
 / FEATURE:
 / MODIFIED TYPE: protein
 / NAME/KEY: Region
 / LOCATION: 1..654
 / OTHER INFORMATION: /note="consensus"
 / US-09-418-540-10

Query Match 25.3%; Score 48.5; DB 4; Length 654;
 Best Local Similarity 25.0%; Pred. No. 50;
 Matches 8; Conservative 12; Mismatches 9; Indels 3; Gaps 1;

QY 3 SYRYKLNNAYQVQVQNKEDAMIEHDVW 31
 DB 348 TYKRGHDYAYKQKTSKTRVPACBRLWK 379

RESULT 9
 / US-09-134-001C-3412
 / Sequence 3412, Application US/09134001C
 / Patent No. 6380370
 / GENERAL INFORMATION:
 / APPLICANT: Lynn Doucette-Stamm et al
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 / INVENTOR: Lynn Doucette-Stamm, Robert G.
 / CURRENT FILING DATE: 1998-08-13
 / PRIOR FILING DATE: 1997-08-14
 / PRIOR APPLICATION NUMBER: US 60/064,964
 / CURRENT FILING DATE: 1998-08-13
 / PRIOR FILING DATE: 1997-08-14
 / PRIOR APPLICATION NUMBER: US 60/055,779
 / NUMBER OF SEQ ID NOS: 5674
 / SEQ ID NO: 178
 / LENGTH: 178
 / TYPE: PRT
 / ORGANISM: Staphylococcus epidermidis
 / US-09-134-001C-3412

Query Match 24.7%; Score 47.5; DB 4; Length 178;
 Best Local Similarity 41.7%; Pred. No. 22;
 Matches 10; Conservative 6; Mismatches 7; Indels 1; Gaps 1;
 QY 1 RSRYRYKLNNAYQVQVQNKEDW 24
 DB 139 RQNHKSQSDMNHQOQHK-DAM 161

RESULT 10
 / US-09-428-517-2
 / Sequence 2, Application US/09428517
 / Patent No. 6251636
 / GENERAL INFORMATION:
 / APPLICANT: Becton, Mary C.
 / APPLICANT: Shah, Sanjay Krishnakant
 / APPLICANT: McDaniel, Robert
 / APPLICANT: Tang, Li
 / TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
 / INVENTOR: Becton, Mary C., Shah, Sanjay Krishnakant, McDaniel, Robert, Tang, Li
 / CURRENT FILING DATE: 1999-10-28
 / EARLIER FILING DATE: 1999-02-16
 / EARLIER APPLICATION NUMBER: 60/120,254
 / EARLIER FILING DATE: 1998-10-29
 / NUMBER OF SEQ ID NOS: 12
 / SEQ ID NO: 2
 / SEQ ID NO: 2
 / SEQ ID NO: 2
 / LENGTH: 4150
 / TYPE: PRT
 / ORGANISM: Artificial Sequence
 / FEATURE:
 / OTHER INFORMATION: Description of Artificial Sequence: Recombinant
 / US-09-428-517-2

Query Match 24.7%; Score 47.5; DB 4; Length 4150;
 Best Local Similarity 44.4%; Pred. No. 7,7e-02;
 Matches 12; Conservative 4; Mismatches 10; Indels 1; Gaps 1;
 QY 5 RYKLNNAYQVQVQNKEDAMIEHDVW 31
 DB 1924 RYKLDQKAKRERIED-WRYRWRK 1949

RESULT 11
 / US-09-121-979-4
 / Sequence 4, Application US/09121979
 / Patent No. 6153709
 / GENERAL INFORMATION:
 / APPLICANT: Holcik, Robert G.
 / APPLICANT: Holcik, Martin
 / TITLE OF INVENTION: XIAP IRES AND USES THEREOF
 / FILE REFERENCE: 07891/021001
 / CURRENT FILING DATE: US/09/121,979
 / PRIOR FILING DATE: 1998-07-24
 / NUMBER OF SEQ ID NOS: 4
 / SEQ ID NO: 1
 / SEQ ID NO: 1
 / SEQ ID NO: 1
 / LENGTH: 236
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 / US-09-121-979-4

Query Match 24.0%; Score 46; DB 4; Length 236;
 Best Local Similarity 39.1%; Pred. No. 50
 Matches 9; Conservative 2; Mismatches 6; Indels 6; Gaps 1;
 QY 8 LNNAYQVQVQNKEDAMIEHDVW 30
 DB 46 LANN-----KPKEDPMDHAWK 62

RESULT 12
 / US-09-332-319-4
 / Sequence 4, Application US/09332319
 / Patent No. 6171821
 / GENERAL INFORMATION:
 / APPLICANT: Korneluk, Robert G.
 / APPLICANT: Holcik, Martin

```

/ APPLICANT: Liatson, Peter
/ TITLE OF INVENTION: XIAP IRS AND USES THEREOF
/ FILE REFERENCE: 07891/021002
/ CURRENT APPLICATION NUMBER: US/09/332,319
/ CURRENT FILING DATE: 1998-07-24
/ EARLIER APPLICATION NUMBER: 09/121,979
/ EARLIER FILING DATE: 1998-07-24
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: FASTSEQ for Windows Version 3.0
/ SEQ ID NO 4
/ LENGTH: 236
/ TYPE: RT
/ ORGANISM: Homo sapiens
US-09-332-319-4

Query Match          24.0%; Score 46; DB 4; Length 236;
Best Local Similarity 39.1%; Pred. No. 50;
Matches          9; Conservative 2; Mismatches 6; Gaps 1;

QY      8 LMANAQQOQKEDAMIEDVW 30
DB      46 LAMN-----KPKEDPMQJAHK 62

RESULT 13
US-09-239-867-2
/ Sequence 2, Application US/09239867,
/ SEQ ID NO 63115
/ GENERAL INFORMATION:
/ APPLICANT: Robert G. Korneluk et al.
/ TITLE OF INVENTION: METHODS AND COMPOUNDS FOR MODULATING
/ FILE REFERENCE: 07891/018002
/ CURRENT APPLICATION NUMBER: US/09/239,867
/ CURRENT FILING DATE: 1998-07-29
/ EARLIER APPLICATION NUMBER: 60/073,001
/ PRIOR FILING DATE: 1998-01-29
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 236
/ TYPE: RT
/ ORGANISM: Homo sapiens
US-09-239-867-2

Query Match          24.0%; Score 46; DB 4; Length 236;
Best Local Similarity 39.1%; Pred. No. 50;
Matches          9; Conservative 2; Mismatches 6; Gaps 1;

QY      8 LMANAQQOQKEDAMIEDVW 30
DB      46 LAMN-----KPKEDPMQJAHK 62

RESULT 14
US-08-746-789A-2
/ Sequence 2, Application US/08746789A
/ SEQ ID NO 5789200
/ APPLICANT: Israel Koli, Martin J. Tymms, Christine Debouck
/ TITLE OF INVENTION: A No. 5789200el Human ETS Family Member, ELP3
/ NUMBER OF SEQUENCES: 4
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: SmithKline Beecham Corporation
/ STREET: 709 Swedeland Road, P.O. Box 1539
/ CITY: King of Prussia
/ STATE: PA
/ COUNTRY: USA
/ ZIP: 19406-0939
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
/ COMPUTER: IBM 486
/ OPERATING SYSTEM: WINDOWS FOR WORKGROUPS

```

```

/ SOFTWARE: MICROSOFT WORD
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/746,789A
/ APPLICATION NO. 789200elmer 15, 1996
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: William T. Han
/ REGISTRATION NUMBER: 34,344
/ REGISTRATION NO.: 3446G 50024
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 610 270 5219
/ TELEFAX: 610 270 4026
/ INFORMATION FOR SEQ ID NO: 2;
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 371
/ TYPE: Amino Acid
/ TOPLOCY: Linear
US-08-746-789A-2

Query Match          23.4%; Score 45; DB 1; Length 371;
Best Local Similarity 47.1%; Pred. No. 1,1e+02;
Matches          8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      7 KLANAAYQOQKEDN 23
DB      71 QVDMISYQVKNKYDA 87

RESULT 15
US-09-342-648-9
/ Sequence 9, Application US/09342648
/ SEQ ID NO 628510
/ GENERAL INFORMATION:
/ APPLICANT: Cahoon, Rebecca E.
/ APPLICANT: Odell, Joan
/ TITLE OF INVENTION: Transcription Coactivators
/ FILE REFERENCE: BB-1169-B
/ CURRENT APPLICATION NUMBER: US/09/342,648
/ CURRENT FILING DATE: 1998-06-29
/ EARLIER APPLICATION NUMBER: 60/092,659
/ EARLIER FILING DATE: July 13, 1998
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 9
/ LENGTH: 885
/ TYPE: RT
/ ORGANISM: Homo sapiens
US-09-342-648-9

Query Match          23.4%; Score 45; DB 4; Length 885;
Best Local Similarity 37.0%; Pred. No. 3e+02;
Matches          4; Mismatches 13; Indels 0; Gaps 0;

QY      2 RSTRYKLANAAYQOQKEDNIEDH 28
DB      611 RBSYVSLSALEAAQKKEKMAHVE 637

Search completed: March 26, 2003, 16:51:40
Job time : 22.7442 secs

```



```

Db 195 RRSRYKLLMAYOOVQONKEDAWIHEHDVWMEI 228
SEQUENCE 338 AA; 39918 MW; ED49E86E067A32B CRC64;

RESULT 2
ID 092450 PRELIMINARY; PRT; 338 AA.
AC 092450
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DB Six transmembrane endothelial antigen of plac.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
NCBI_TaxID=96823;
RN (1)
SEQUENCE FROM N.A.
RA Nagaoka T., Boudry G., Coupel S., Coulon F., Tesson L.,
RA Haslan U.-M., Souillion U.-P., Charreau B.1 cells during TNF-alpha-
RA and LPS-mediated activation". In endothelial cells during TNF-alpha-
RA and LPS-mediated activation".
DR Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF319659; AAC33868.1; -.
KW Transmembrane.
SQ
SEQUENCE 338 AA; 39918 MW; ED49E86E067A32B CRC64;

Query Match 100.0%; Score 197; DB 5; Length 338;
Spec. Local Similarity 100.0%; Tred No. 176-177;
Matched 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RRSRYKLLMAYOOVQONKEDAWIHEHDVWMEI 34
Db 184 RRSRYKLLMAYOOVQONKEDAWIHEHDVWMEI 217

RESULT 3
ID 092450 PRELIMINARY; PRT; 339 AA.
AC 092450
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DB 24103981X
GN STRAP OR 24100075191X.
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN (1)
SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J TISSUE=EMBRYONIC STEM CELLS;
RC MEDLINE=21085650; Pubmed=1121781;
RX Kawai U., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa T., Hara A., Fukunishi Y., Kono H., Aoechi Y., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Kiehlmann M., Gaudet J., Hoshida S., Kage H., G. Quecunhu J.,
RA Kiehlmann M., Gaudet J., Hoshida S., Kage H., G. Quecunhu J.,
RA Schirral L.M., Strubli P., Suzuki R., Tomita M., Wagner L., Maehiro T.,
RA Sakai K., Okido T., Furuno M., Kono H., Balderelli R., Baran G.,
RA Blake J., Boffelli D., Bojunga N., Carucci P., de Bonaldo M.F.,
RA Guelinche M.J., Hall D., Hofman M., Huie D.A., Kamiya M., Lee N.H.,
RA Lyons P., Matchioni L., Mahina J., Mazzarelli D., Nomaets P.,
RA Norberg B., Salong B., Schlegel M., Rodriguez S., Sakamoto S.,
RA Norberg B., Salong B., Schlegel M., Rodriguez S., Sakamoto S.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Wittkecker C., Wilmig L.,
RA Wyszewski A., Yoshida K., Hasegawa Y., Kawai H., Kohzuki S.,
RA Hayashiaki Y.
RT "Functional annotation of a full-length mouse cDNA collection."
DR Nature 409:685-690(2001).
RL EMBL; AK010437; BAB2638.1; -.

```

```

DR MGD; MG1:1917608; Strap.
SQ
SEQUENCE 339 AA; 39264 MW; 3F7AB9C7520F0966 CRC64;

Query Match 97.4%; Score 187; DB 11; Length 339;
Spec. Local Similarity 94.1%; Tred No. 1,6e-11;
Matched 32; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RRSRYKLLMAYOOVQONKEDAWIHEHDVWMEI 34
Db 185 RRSRYKLLMAYOOVQONKEDAWIHEHDVWMEI 218

RESULT 4
ID 092459 PRELIMINARY; PRT; 339 AA.
AC 092459
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DB Six transmembrane epithelial antigen of the prostate.
GN STRAP
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN (1)
SEQUENCE FROM N.A.
RA Yang D., Holt G.E., Veldeers M.P., Kwon E.D., Kaet W.M.;
RA Yang D., Holt G.E., Veldeers M.P., Kwon E.D., Kaet W.M.;
RT "Murine six-transmembrane epithelial antigen of the prostate, prostate
RT stem cell antigen, and prostate-specific membrane antigen: prostate
RT specific cell-surface antigen highly expressed in prostate cancer of
RT transgenic adenocarcinoma of the prostate mice."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF297098; AAK8126.1; -.
KW MGD; MG1:1917608; Strap.
SQ
SEQUENCE 339 AA; 4826471F559894F CRC64;

Query Match 97.4%; Score 187; DB 11; Length 339;
Spec. Local Similarity 94.1%; Tred No. 1,6e-11;
Matched 32; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RRSRYKLLMAYOOVQONKEDAWIHEHDVWMEI 34
Db 185 RRSRYKLLMAYOOVQONKEDAWIHEHDVWMEI 218

RESULT 5
ID 092422 PRELIMINARY; PRT; 339 AA.
AC 092422
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DB Ductal in situ carcinoma of the prostate.
GN Ductal in situ carcinoma of the prostate.
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN (1)
SEQUENCE FROM N.A.
RA Serra V., Mavret P., Lambin D., Vaudrouelle M., Kellermann O.,
RA Serra V., Mavret P., Lambin D., Vaudrouelle M., Kellermann O.,
RT "Prostate and non-prostate expression of ductin, the mouse ortholog
RT of human STRAP."
DR Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY029584; AAK50537.1; -.
KW MGD; MG1:1915678; 1010001D01R1X.
SQ
SEQUENCE 339 AA; 39109 MW; 32A2C29F2E33B0D CRC64;

```


Best Local Similarity 55.9%; Pred. No. 0.00014;
 Matches 19; Conservative 6; Mismatches 8; Indels 1; Gaps 1;
 QY 1 RRSRYKLLMAYQVQONKEDAMIEHDYMEI 34
 DB 326 RRAHRIIDVNLVAVQVLANSLSHLMVE-EWAMEI 358

RESULT 10

Q08703 PRELIMINARY; PRT; 487 AA.
 ID Q08703
 AC Q08703
 DT 01-JUN-2002 (T-EMBLrel. 21, Created)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE Tumor suppressor p14de.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN [1]_TaxId=9606;
 RP SEQUENCE FROM N.A.
 RA Steiner M.S., Allay J.A., Wang C.;
 RT "A novel prostate-derived tumor suppressor."
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
 DB EMBL; AF238864; ALN18206.1;
 SQ SEQUENCE 487 AA; 54402 MW; AFT6053590E6F68 CRC64;

Query Match 50.3%; Score 96.5; DB 4; Length 487;
 Best Local Similarity 55.9%; Pred. No. 0.00015;
 Matches 19; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 1 RRSRYKLLMAYQVQONKEDAMIEHDYMEI 34
 DB 326 RRAHRIIDVNLVAVQVLANSLSHLMVE-EWAMEI 358

RESULT 11

067595 PRELIMINARY; PRT; 480 AA.
 ID 067595
 AC 067595
 DT 01-AUG-1998 (T-EMBLrel. 07, Created)
 DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE Hypothetical protein A0_1687.
 GN A0_1687.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OC NCBI_TaxId=63163;
 RN [1]_TaxId=63163;
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98196666; PubMed=9537220;
 RA Deckert G., Warren P.V., Gaasbeek R., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Speed M.A., Keller M., Anjay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.V., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus." 92:353-358(1998).
 DB EMBL; A800751; AAC07560.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 480 AA; 55582 MW; 7BBD89517A218046 CRC64;

Query Match 32.6%; Score 62.5; DB 16; Length 480;
 Best Local Similarity 25.0%; Pred. No. 4;
 Matches 12; Conservative 17; Mismatches 4; Indels 15; Gaps 2;

QY 1 RRSRYKLLMAYQVQONKEDAMIEHDYMEI 33
 DB 10 RRTYRIKIPAKVSIQCKEYKVLMSYSGRIEKSKEDEVKQKYAVK 57

RESULT 12

Q08749 PRELIMINARY; PRT; 926 AA.
 ID Q08749
 AC Q08749
 DT 01-JUN-2002 (T-EMBLrel. 21, Created)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE Cezanne 2 protein.
 GN CEZANNE 2.
 OS Homo sapiens (human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN [1]_TaxId=9606;
 RP SEQUENCE FROM N.A.
 RA Evans P.C., Coadwell W.J., Klishaw P.J.;
 RT "Isolation of a novel human gene, Cezanne 2,"
 RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
 DB EMBL; AF430383; C002107.1;
 SQ SEQUENCE 926 AA; 100676 MW; 6846232CB2C8058 CRC64;

Query Match 31.2%; Score 60; DB 4; Length 926;
 Best Local Similarity 64.3%; Pred. No. 17;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 11 WAYQVQONKEDAM 24
 DB 250 WNRQVQONKEDAM 263

RESULT 13

Q08554 PRELIMINARY; PRT; 926 AA.
 ID Q08554
 AC Q08554
 DT 01-JUN-2002 (T-EMBLrel. 21, Created)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE Cezanne 2 protein.
 GN CEZANNE 2.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]_TaxId=10090;
 RP SEQUENCE FROM N.A.
 RA STRAIN=C57 BL/6;
 RA Evans P.C., Coadwell W.J., Klishaw P.J.;
 RT "Isolation of a novel murine gene, Cezanne 2,"
 RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
 DB EMBL; AF430384; C002108.1;
 SQ SEQUENCE 926 AA; 100796 MW; 4B6BD05A0108BD9 CRC64;

Query Match 31.2%; Score 60; DB 11; Length 926;
 Best Local Similarity 64.3%; Pred. No. 17;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 11 WAYQVQONKEDAM 24
 DB 252 WNRQVQONKEDAM 265

RESULT 14

Q09507 PRELIMINARY; PRT; 1027 AA.
 ID Q09507
 AC Q09507
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE Putative TRP homologous cation channel protein (fragment).
 GN TRPC4/CE0-11.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditia.
 OC NCBI_TaxId=6239;

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using SW model

Run on: March 26, 2003, 16:38:14 ; Search time 66.0233 Seconds
(without alignments)
68.620 Million cell updates

Title: US-10-010-667A-20
Perfect score: 100

Sequence: 1 RRSYRYKLNTWAYQQVQNKEDAWIEHDVWRMEI 34

Scoring table: BLOSUM62

Gapor 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Maximum Match 100%

Listing first 45 summaries

Database :

1	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1980.DAT*
2	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1981.DAT*
3	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1982.DAT*
4	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1983.DAT*
5	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1984.DAT*
6	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1985.DAT*
7	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1986.DAT*
8	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1987.DAT*
9	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1988.DAT*
10	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1989.DAT*
11	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1990.DAT*
12	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1991.DAT*
13	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1992.DAT*
14	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1993.DAT*
15	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1994.DAT*
16	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1995.DAT*
17	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1996.DAT*
18	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1997.DAT*
19	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1998.DAT*
20	S1D52.gcdgata.geneseq.geneseq-emb1.AAI1999.DAT*
21	S1D52.gcdgata.geneseq.geneseq-emb1.AAI2000.DAT*
22	S1D52.gcdgata.geneseq.geneseq-emb1.AAI2001.DAT*
23	S1D52.gcdgata.geneseq.geneseq-emb1.AAI2002.DAT*

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	192	100.0	34	21	AA58189	Human STRP-1 protein
2	192	100.0	34	21	AA582187	Extracellular loop 1 of
3	192	100.0	34	22	AA582187	Extracellular loop 1 of
4	192	100.0	129	22	AA875315	Human secreted protein
5	192	100.0	339	21	AA581894	Human STRP-1 protein
6	192	100.0	339	21	AA569927	Human prostate CDN
7	192	100.0	339	22	AA478845	Human protein SBO
8	192	100.0	339	22	AA401282	P7899 amino acid
9	192	100.0	339	23	AB64132	Human prostate cancer
10	192	100.0	339	23	AB645817	Human P7899 protein

45	52	27.1	38	23	AB041121
44	52	27.1	38	22	AA011599
43	52	27.1	38	22	AA011991
42	52	27.1	38	22	AA017318
41	52	27.1	38	22	AB058802
40	52	27.1	38	22	AB032132
39	52	27.1	38	22	AB031059
38	52	27.1	38	22	AB031059
37	52	27.1	38	22	AB031059
36	52	27.1	38	22	AB031059
35	52	27.1	38	22	AB031059
34	52	27.1	38	22	AB031059
33	52	27.1	38	22	AB031059
32	52	27.1	38	22	AB031059
31	52	27.1	38	22	AB031059
30	52	27.1	38	22	AB031059
29	52	27.1	38	22	AB031059
28	52	27.1	38	22	AB031059
27	52	27.1	38	22	AB031059
26	52	27.1	38	22	AB031059
25	52	27.1	38	22	AB031059
24	52	27.1	38	22	AB031059
23	52	27.1	38	22	AB031059
22	52	27.1	38	22	AB031059
21	52	27.1	38	22	AB031059
20	52	27.1	38	22	AB031059
19	52	27.1	38	22	AB031059
18	52	27.1	38	22	AB031059
17	52	27.1	38	22	AB031059
16	52	27.1	38	22	AB031059
15	52	27.1	38	22	AB031059
14	52	27.1	38	22	AB031059
13	52	27.1	38	22	AB031059
12	52	27.1	38	22	AB031059
11	52	27.1	38	22	AB031059
10	52	27.1	38	22	AB031059
9	52	27.1	38	22	AB031059
8	52	27.1	38	22	AB031059
7	52	27.1	38	22	AB031059
6	52	27.1	38	22	AB031059
5	52	27.1	38	22	AB031059
4	52	27.1	38	22	AB031059
3	52	27.1	38	22	AB031059
2	52	27.1	38	22	AB031059
1	52	27.1	38	22	AB031059

ALIGNMENTS

XX	Result 1
XX	AAVS6199
XX	AAVS8199 standard; peptide: 34 AA.
XX	AAVS8199;
XX	14-MAR-2000 (first entry)
XX	Human STRAP-1 peptide, corresponding to STRAP-1 extracellular region 2
XX	Serputive transmembrane antigen of the prostate; STRAP-1; prostate;
XX	transmembrane domain; type IIIa membrane protein; expression; cancer;
XX	prostate cancer; bladder cancer; colon cancer; pancreatic cancer;
XX	ovarian cancer; tumour antigen; immunisation; immune response;
XX	cellular; humoral; anticancer vaccine; antibody; detection; diagnosis;
XX	prognosis; monitoring; susceptibility; therapeutic inhibitor;
XX	drug targeting; recombinant protein.
XX	Synthetic.
XX	Homo sapiens.
XX	MO962941-A2.
XX	09-DEC-1999.
XX	01-JUN-1999; 99MO-US12157.
XX	01-JUN-1998; 98US-0087520.
XX	30-JUN-1998; 98US-0091183.
XX	(UROG)- UROGNESSYS INC.
XX	(AFAR) AFAR D B.
XX	(HUSE) HUBERT R S.

PA	(LEON)/ LEONG K.
PA	(RAIT)/ RAITANO A. B.
PA	(SAFF)/ SAFFRAN D. C.
PA	
XX	Afar DE, Hubert RS, Leong K, Raitano AB, Saffran DC,
DR	WPIJ, 2000-072832/06.
PT	
PT	Novel proteins useful as diagnostic markers and therapeutic targets,
PT	particularly for prostatic cancer -
XX	
XX	Disclosure, Page 22; 83pp. English.
CC	Sequences AA159198-Y58200 represent synthetic peptides that correspond
CC	to the extracellular regions of STRAP-1 (serpinine transmembrane
CC	antigen of the prostate, AA159194). These peptides were used to raise
CC	monoclonal anti-STRAP-1 antibodies. STRAP-1 is the prototype
CC	member of the STRAP family of proteins (AA159194-Y58197) which
CC	exhibit a high degree of structural conservation, but which show
CC	no significant structural homology to known human proteins. The STRAP-1
CC	gene is a member of the serpinin gene family and encodes a type I
CC	single ITD membrane protein and is expressed predominantly in prostatic
CC	cells in normal human tissues. Structurally, STRAP-1 is a 339 amino
CC	acid protein characterized by six transmembrane domain and a
CC	intracellular N- and C-terminal, suggesting that it folds in a
CC	"serpinine" manner into three extracellular and two intracellular loops.
CC	STRAP-1 mRNA and protein expression is maintained at high levels and
CC	throughout all stages of prostate cancer. STRAP-1 is a secreted protein
CC	and its expression is up-regulated in metastatic prostate cancer. It is
CC	pancreatic and ovarian cancer. The function of the STRAP protein is not
CC	known. They may be ion channels (from the presence of six transmembrane
CC	domains), a feature which is shared by certain ion channels) or
CC	GAP-junction proteins (from immunohistochemical staining). STRAP-1 and
CC	STRAP-2 are cell-surface tumour antigens. Immunisation with a STRAP
CC	protein induces cellular and humoral immune responses against
CC	prostate cancer. Strong evidence suggests that STRAP-1 is involved in
CC	specific antibody agents. To produce anticancer vaccines and to generate
CC	specific antibodies. The antibodies may be used for detection, prognostic,
CC	and monitoring of cancers (or susceptibility to cancer), as therapeutic
CC	inhibitors or to target therapeutic agents to their site of action. STRAP
CC	nucleotide acids may be used for recombinant protein production, as
CC	diagnostic and prognostic reagents, for identifying STRAP-expressing
CC	cells for screening inhibitors of STRAP expression and for therapeutic
CC	modulation/inhibition of protein expression. This is usually targeted by
CC	systemically administered agents, and because they are expressed mainly
CC	on prostatic epithelial cells, agents targeted to them should have
CC	minimal side effects on other tissues.
XX	
SX	Sequence 34 AA:
DB	
QY	1 RRSRYKLMMAYQQVQNKKEDMATHENDYVMNKE 34
DB	1 RRSRYKLMMAYQQVQNKKEDMATHENDYVMNKE 34
RESUOT 2	
ID	AAE02787 standard; peptide; 34 AA.
AC	AAE02787
AC	AAE02787;
DT	06-AUG-2001 (first entry)
XX	
XX	Extracellular loop #2 of human STRAP-1, suitable for cloning into pGc.
KM	Human; cytosolic; antiproliferative; vaccine; gene therapy;
KM	six transmembrane epithelial antigen of the prostate-1; STRAP-1;
KM	chromosome 7p22.3; cancer; prostate; colon; bladder; pancreatic;

XX	ovarian, lung; extracellular loop; serpentine transmembrane antigen.
XX	
XX	Homo sapiens.
XX	
XX	MO2001.40276-A2.
XX	
XX	07-JUN-2001.
XX	
XX	06-DEC-2000; 2000MO-W531040.
XX	
XX	06-DEC-1999; 9905-0455486.
XX	
XX	(UROC-) UROGENESIS INC.
XX	
XX	Afar DEH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC, Faris M; Jakobovits A;
XX	
XX	WFI; 2001-367804/38.
XX	
XX	Now STEAP (six transmembrane epithelial antigen of the prostate) protein, expressed in human cancers, useful for detecting and treating cancer -
XX	
XX	
XX	Example 19; Page 102; 187p; English.
XX	
XX	The present invention relates to human six transmembrane epithelial antigen of the prostate (STEAP) proteins. STEAP proteins are used in gene therapy, inhibiting the development or progression of a cancer (eg. prostate, colon, bladder, lung, ovarian and pancreatic) expressing STEAP or inhibiting growth or killing cells expressing STEAP in a patient. completes administering a vaccine composition to the patient. Treating a patient with a cancer that expresses STEAP, or inhibiting growth or killing cells expressing STEAP, completes administering to a patient a vaccine composition of the heavy and light chains of the monoclonal antibody that specifically binds to STEAP, such that the vector delivers the single chain monoclonal antibody coding sequence to the cancer cells and the encoded single chain monoclonal antibody is expressed intracellularly. The present sequence is extracellular loop of STEAP-1 suitable for cloning into pFC, which is used in the invention. STEAP-1 gene is located on chromosome 1p22.3.
XX	
XX	
XX	Sequence 34 Ab.
XX	
XX	Query Match 100.0%; Score 192; DB 22; Length 34;
XX	Best Local Similarity 100.0%; Pred. NO. 3.7e-19;
XX	Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX	
XX	1 RRSRIKLINAAVQVQCNEDNINIEHDVWKEI 34
XX	
XX	1 RRSRIKLINAAVQVQCNEDNINIEHDVWKEI 34
XX	
XX	1 RRSRIKLINAAVQVQCNEDNINIEHDVWKEI 34
XX	
XX	1 RRSRIKLINAAVQVQCNEDNINIEHDVWKEI 34
XX	
XX	AAAT75314.
XX	
XX	03-APR-2001 (first entry)
XX	
XX	Gene 20 human secreted protein homologous amino acid sequence #133.
XX	
XX	Human; immunosuppressive; antiarthritic; antineumatic; noctropic; antiapoptotic; cytotoxic; candida; vaxodotropic; cardioprotective; neuroprotective; antibacterial; vasoprotective; neuroprotective; cardiovascular disorder; cerebrovascular disorder; infection; nervous system disorder; ocular disorder; chemotaxis; food additive; secreted protein.
XX	
XX	Homo sapiens.
XX	

XX MO20007021-A1.
 XX 21-DEC-2000.
 XX 01-JUN-2000; 2000MO-US15135.
 XX 11-JUN-1999; 99US-0138632.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX (ROSE/) ROSEN C A.
 XX Rosen CA, Ruben SM, Komatsoulis GA;
 XX WPI; 2001-071257/08.
 XX Nucleic acid molecules encoding human secreted proteins, used in
 XX preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
 XX Parkinson's diseases and cancer -
 XX Disclosure: Page 53-54; 530pp; English.

XX This invention relates to polynucleotide sequences AAF63789 - AAF63836
 XX which encode human secreted proteins AAB75260 - AAB75287. Included in the
 XX invention are protein sequences AAB75288 - AAB75341 which are fragments
 XX of the secreted proteins and amino acid sequences with which these
 XX fragments share homology. Examples of the activities of the proteins and
 XX polynucleotides and the activities of the amino acid sequences and
 XX fragments are: immunosuppressive; antiallergic; antithrombotic; antineoplastic;
 XX antiproliferative; cytostatic; cardiatic; vasotropic; cerebroprotective;
 XX neurotropic; neuroprotective; and vulnary activity. The protein and polynucleotide
 XX sequences, their agonists and antagonists may be useful for treating,
 XX preventing and diagnosing diseases and disorders such as autoimmune
 XX diseases e.g. rheumatoid arthritis, hyperproliferative disorders
 XX e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia,
 XX angioneuroma, nervous system disorders e.g. Alzheimer's disease,
 XX infections caused by bacteria, viruses and fungi and ocular disorders
 XX e.g. corneal infection. The polypeptides can also be used to aid wound
 XX healing and epithelial cell proliferation, to prevent skin aging due to
 XX sunburn, to maintain organs before transplantation, for supporting cell
 XX culture of primary tissues, to regenerate tissues and in chemotaxis. The
 XX polypeptides can also be used as adjuvants in vaccines. The amino acid
 XX sequences and fragments and amino acid sequences included in the invention are
 XX oligonucleotides AAF63780 - AAF63788 and peptide AAB75239 which are used
 XX in the identification and characterisation of the DNA and protein
 XX sequences of the invention.

XX Sequence 129 AA;
 XX
 XX Query Match 100.0%; Score 192; DB 22; Length 129;
 XX Best Local Similarity 100.0%; Pctd No. 1, 7e-18;
 XX Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLMAYQVQVQNKEDMTEHDYVMEI 34
 DB 60 RRSRYKLLMAYQVQVQNKEDMTEHDYVMEI 93

RESULT 4
 AAB75315
 ID AAB75315 standard; Protein: 129 AA.
 XX AAB75315;
 XX
 XX 03-APR-2001 (first entry)
 XX
 XX Human secreted protein sequence encoded by gene 20 SEQ ID NO:134.
 XX
 XX Human; immunosuppressive; antiallergic; antithrombotic; neurotropic;
 XX antiproliferative; cytostatic; cardiatic; vasotropic; cerebroprotective;
 XX neuroprotective; antibacterial; virucide; fungicide; ophthalmological;

KW vulnary; autoimmune disease; hyperproliferative disorder; cancer;
 KW cardiovascular disorder; cerebrovascular disorder; infection;
 KW nervous system disorder; ocular disorder; chemotaxis; food additive;
 KW secreted protein.
 OS Homo sapiens.
 PN MO20007021-A1.
 XX 21-DEC-2000.
 XX 01-JUN-2000; 2000MO-US15135.
 XX 11-JUN-1999; 99US-0138632.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX (ROSE/) ROSEN C A.
 XX Rosen CA, Ruben SM, Komatsoulis GA;
 XX WPI; 2001-071257/08.
 XX Nucleic acid molecules encoding human secreted proteins, used in
 XX preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
 XX Parkinson's diseases and cancer -
 XX Disclosure: Page 53-54; 530pp; English.

XX This invention relates to polynucleotide sequences AAF63789 - AAF63836
 XX which encode human secreted proteins AAB75260 - AAB75287. Included in the
 XX invention are protein sequences AAB75288 - AAB75341 which are fragments
 XX of the secreted proteins and amino acid sequences with which these
 XX fragments share homology. Examples of the activities of the proteins and
 XX polynucleotides and the activities of the amino acid sequences and
 XX fragments are: immunosuppressive; antiallergic; antithrombotic; antineoplastic;
 XX antiproliferative; cytostatic; cardiatic; vasotropic; cerebroprotective;
 XX neurotropic; neuroprotective; and vulnary activity. The protein and polynucleotide
 XX sequences, their agonists and antagonists may be useful for treating,
 XX preventing and diagnosing diseases and disorders such as autoimmune
 XX diseases e.g. rheumatoid arthritis, hyperproliferative disorders
 XX e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia,
 XX angioneuroma, nervous system disorders e.g. Alzheimer's disease,
 XX infections caused by bacteria, viruses and fungi and ocular disorders
 XX e.g. corneal infection. The polypeptides can also be used to aid wound
 XX healing and epithelial cell proliferation, to prevent skin aging due to
 XX sunburn, to maintain organs before transplantation, for supporting cell
 XX culture of primary tissues, to regenerate tissues and in chemotaxis. The
 XX polypeptides can also be used as adjuvants in vaccines. The amino acid
 XX sequences and fragments and amino acid sequences included in the invention are
 XX oligonucleotides AAF63780 - AAF63788 and peptide AAB75239 which are used
 XX in the identification and characterisation of the DNA and protein
 XX sequences of the invention.

XX Sequence 129 AA;
 XX
 XX Query Match 100.0%; Score 192; DB 22; Length 129;
 XX Best Local Similarity 100.0%; Pctd No. 1, 7e-18;
 XX Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRSRYKLLMAYQVQVQNKEDMTEHDYVMEI 34
 DB 60 RRSRYKLLMAYQVQVQNKEDMTEHDYVMEI 93

RESULT 5
 AAY58194
 ID AAY58194 standard; Protein: 339 AA.
 XX AAY58194;
 XX

DT 14-MAR-2000 (first entry)
 XX Human STRAP-1 protein.
 XX Serpentine transmembrane antigen of the prostate; STRAP-1; prostate;
 XX antigen that binds the 11L membrane protein; expression cancer;
 XX prostate cancer; bladder cancer; colon cancer; pancreatic cancer;
 XX ovarian cancer; tumour antigen; immunisation; immune response;
 XX cellular; humoral; anticancer vaccine; antibody; detection; diagnosis;
 XX prognosis; monitoring; susceptibility; therapeutic inhibitor;
 XX drug targeting; recombinant protein.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 1..69
 FT /note= "Intracellular region 1"
 FT Domain 70..91
 FT /note= "Transmembrane domain 1"
 FT Region 92..113
 FT /note= "Extracellular region 1 (AAV58198)"
 FT Domain 114..136
 FT /note= "Transmembrane domain 2"
 FT Region 137..162
 FT /note= "Intracellular region 2"
 FT Domain 163..184
 FT /note= "Transmembrane domain 3"
 FT Region 185..218
 FT /note= "Extracellular region 2 (AAV58199)"
 FT Domain 219..241
 FT /note= "Transmembrane domain 4"
 FT Region 242..253
 FT /note= "Intracellular region 3"
 FT Domain 252..276
 FT /note= "Transmembrane domain 5"
 FT Region 277..291
 FT /note= "Extracellular region 3 (AAV58200)"
 FT Domain 292..313
 FT /note= "Transmembrane domain 6"
 FT Region 314..339
 FT /note= "Intracellular region 4"
 XX
 XX MO9962941-A2.
 XX
 PD 09-DEC-1999.
 XX
 PF 01-JUN-1999; 99MO-US12157.
 XX
 XX 01-JUN-1998; 98US-0087520.
 PR 30-JUN-1998; 98US-0091183.
 XX
 XX (URC-) UROGENESIS INC.
 PA (AFAR/) AFAR D E.
 PA (HUBE/) HUBERT R. S.
 PA (LEON/) LEONG K.
 PA (RAIT/) RAITANO A B.
 PA (SAFF/) SAFFRAN D C.
 XX
 XX Afar DE, Hubert RS, Leong K, Raitano AB, Saffran DC;
 DR WPI: 2000-072832/06.
 DR N-PSDB; AA249355, AA249396.
 PT Novel proteins useful as diagnostic markers and therapeutic targets,
 PT particularly for prostate cancer
 XX
 XX Claim 1; Fig 1a; 83pp. English.
 CC This sequence represents a novel human protein, STRAP-1 (serpentine
 CC transmembrane antigen of the prostate). STRAP-1 is the prototype
 CC member of the STRAP family of proteins (AAV58194-Y58197) which
 CC exhibit a high degree of structural conservation, but which show
 CC no significant structural homology to known human proteins. The STRAP-1

CC gene has been localised to chromosome 7p22. STRAP-1 is thought to be a
 CC type IIIa membrane protein and is expressed predominantly in prostate
 CC cells in normal human tissues. Structurally, STRAP-1 is a 339 amino
 CC acid protein characterised by a signal sequence, a hydrophobic
 CC region, and a cytoplasmic tail suggesting that it folds in a
 CC "serpentine" manner into three extracellular and two intracellular loops.
 CC STRAP-1 mRNA and protein expression is maintained at high levels and
 CC is also overexpressed in certain other cancers, including bladder, colon,
 CC pancreatic and ovarian cancer. The function of the STRAP proteins is not
 CC known. They may be ion channels, as from the presence of transmembrane
 CC domains, or they may be receptors, as from the presence of extracellular
 CC loop-junction proteins (from immunohistochemical staining). STRAP-1 and
 CC STRAP-2 are cell-surface tumour antigens. Immunisation with a STRAP
 CC protein induces cellular and humoral immune responses against
 CC STRAP-expressing cells. STRAP proteins may be used to identify
 CC specific-binding agents, to produce anticancer vaccines and to generate
 CC specific antibodies. The antibodies may be used for detection, prognosis,
 CC and monitoring of cancers (or susceptible to their effects of action). STRAP
 CC nucleic acids may be used for recombinant protein production, as
 CC diagnostic and prognostic reagents, for identifying STRAP-expressing
 CC cells for screening inhibitors of STRAP expression and for therapeutic
 CC modulation/inhibition of STRAP expression. Since high levels of STRAP
 CC proteins are exposed on the cell surface, they are easily targeted by
 CC systemically administered agents, and because they are expressed mainly
 CC on prostatic epithelial cells, agents targeted to them should have
 CC minimal side effects on other tissues.
 XX
 SQ Sequence 339 AA:
 Query Match 100.0%; Score 192; DB 21; Length 339;
 Best Local Similarity 100.0%; Pred. No. 5; 1e-18;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRSRYKLLMAVYQVQNKEDWIEHDVWMEI 34
 DB 185 RRSRYKLLMAVYQVQNKEDWIEHDVWMEI 218
 |||
 RESULT 6
 AA069927
 MOU69927 standard; Protein; 339 AA.
 AC AA069927;
 XX
 DT 30-JAN-2002 (first entry)
 XX
 DE Human prostate cDNA encoded protein #12.
 XX
 XX Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.
 XX
 OS Homo sapiens.
 XX
 XX MO200171032-A2.
 PD 04-OCT-2001.
 XX
 XX 27-MAR-2001; 2001MO-US09919.
 XX
 XX 27-MAR-2000; 2000US-0536857.
 PR 09-MAY-2000; 2000US-0568100.
 PR 12-MAY-2000; 2000US-0570737.
 PR 13-JUN-2000; 2000US-0593793.
 PR 27-JUN-2000; 2000US-0605783.
 PR 28-AUG-2000; 2000US-0631212.
 PR 28-AUG-2000; 2000US-0651279.
 PR 06-SEP-2000; 2000US-0657229.
 PR 02-OCT-2000; 2000US-0679466.
 PR 10-OCT-2000; 2000US-0685166.
 XX
 XX (CONF-) CORIXA CORP.
 PA
 XX

PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 PI Fanger GR, Retter WM, Stolk JA, Day CH, Vedvick TS, Carter D;
 PI Li SX, Wang A, Skeiky IAW, Hepler WT, Henderson RA;
 DR N-PSDB; AAK64160.
 WPI: 2001-639232/73.
 XX
 PT New human prostate-specific polypeptides and polynucleotides useful for
 the diagnosis and treatment of cancer, especially prostate cancer -
 XX
 XX Claim 2, Page 549; 579pp; English.
 CC The invention relates to isolated prostate-specific
 CC polynucleotides, polypeptides, fusion proteins of the polypeptides,
 CC antibodies raised against the polypeptides (or antigenic epitopes
 CC derived from them) and antigen-presenting cells expressing the
 CC polypeptides. The antibodies are useful for detecting the presence of
 CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
 CC the antigen-presenting cells are useful for stimulating and/or expanding
 CC of cancer especially prostate cancer. Compositions comprising the
 CC polynucleotide and/or polypeptide are useful for stimulating an immune
 CC response, and for treating cancer. The oligonucleotide is useful for
 CC detecting cancer. The present sequence is a prostate specific
 CC polypeptide of the invention.
 XX
 SQ Sequence 339 AA;
 Query March 100.0%; Score 192; DB 22; Length 339;
 Best Local Similarity 100.0%; Pred. No. 5, 1e-18;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRSRYKILMAVYQVQONKEDAWIHDVWMEI 34
 DB 185 RRSIRIKLMAVYQVQONKEDAWIHDVWMEI 218
 RESULT 7
 AAM78845
 ID AAM78845 standard; Protein; 339 AA.
 AC AAM78845;
 XX
 XX 06-NOV-2001 (first entry)
 DE Human protein SEQ ID NO 1507.
 XX
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 XX vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukemia;
 XX nervous system disorder; arthritis; inflammation.
 OS Homo sapiens.
 XX
 XX MO200157139-A2.
 PI
 XX
 PD 09-AUG-2001.
 XX
 XX 05-FEB-2001; 2001MO-US04098.
 XX
 XX 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUL-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663551.
 PR 01-OCT-2000; 2000US-0681351.
 PR 30-NOV-2000; 2000US-0728432.
 XX
 XX (HYSE-) HYSED INC.
 XX
 XX Tang YT, Liu C, Ormanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y,
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejman T, Goodrich R;
 XX
 XX WPI: 2001-476283/51.
 DR N-PSDB; AAK51978.
 XX
 PT Nucleic acid encoding polypeptides with cytokine-like activities,
 PT use in diagnosis and gene therapy -
 XX
 XX Claim 20; Page 3800-3801; 6221pp; English.
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78132-AAK80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g., stem cell growth factor activity, hematopoietic regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC antitumor activity and may be useful in diagnosis, immunotherapy and
 CC inflammation, cancer, leukemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAK80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX
 SQ Sequence 339 AA;
 Query March 100.0%; Score 192; DB 22; Length 339;
 Best Local Similarity 100.0%; Pred. No. 5, 1e-18;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 RRSRYKILMAVYQVQONKEDAWIHDVWMEI 34
 DB 185 RRSIRIKLMAVYQVQONKEDAWIHDVWMEI 218
 RESULT 8
 AAM01282
 ID AAM01282 standard; Protein; 339 AA.
 AC AAM01282;
 XX
 XX 04-OCT-2001 (first entry)
 DE P789P amino acid sequence.
 XX
 XX Human; prostate cancer; prostate-specific; diagnosis; vaccine;
 KW cytotactic; gene therapy; metastasis.
 OS Homo sapiens.
 XX
 XX MO200151633-A2.
 PI
 XX
 PD 19-JUL-2001.
 XX
 XX 16-JAN-2001; 2001MO-US01574.
 XX
 XX 14-JAN-2000; 2000US-0483672.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG,
 PI Kalos MD, Fanger GR, Day CH, Retter WM, Stolk JA, Skeiky IAW;
 PI Wang A, Weagher MJ;
 DR WPI: 2001-425873/45.
 XX
 PT New polynucleotide encoding a prostate-specific protein, for
 PT diagnosing, monitoring and treating prostate cancer in a patient and
 PT for use in vaccines -
 XX
 XX Claim 2, Page 510-512; 543pp; English.

XX Kidney injury associated molecule HM018 protein.
 XX Kidney injury associated molecule, kidney injury related molecule;
 XX KIM, tissue growth promotion; regeneration; renal condition;
 XX acute renal failure; acute nephritis; tumour.
 OS Rattus sp.
 XX MO9853071-N1.
 XX 26-NOV-1998.
 XX 22-MAY-1998; 98NO-US10547.
 XX 23-MAY-1997; 97US-0047491.
 XX 23-MAY-1997; 97US-0047490.
 XX (BIO) BIOGEN INC.
 PI Cate RL, Hession CM, Santicola-Nadel M, Wei H;
 DR NPI; 1999-045312/04.
 DR N-PSDB; AAB60586.
 XX Kidney injury associated molecule, KIM, polypeptides - upregulated
 PT in injured/regenerating kidney, renal tissue growth
 PT and regeneration, especially to treat renal conditions
 PS Claim 17; Page 57-58; 213pp; English.
 XX The present sequence represents a kidney injury associated molecule
 CC (KIM) protein. KIM proteins can be administered therapeutically
 CC by expressing KIM encoding polynucleotides, to promote growth and/or
 CC by administering KIM proteins, to promote growth and/or
 CC KIM proteins can be used to regenerate (especially renal) tissues
 CC KIM fusion proteins, conjugates, antibodies and vectors can also be used
 CC therapeutically, e.g. these or the KIM proteins may be included with an
 CC acceptable carrier in pharmaceutical compositions, useful for therapy/
 CC prophylaxis of conditions associated with dysfunction/dysregulation of
 CC KIM genes or proteins, especially renal diseases or impairments of renal
 CC function. Examples e.g. related to renal diseases include: renal
 CC dysfunction, renal failure, renal insufficiency, renal impairment, renal
 CC internalised into cells, can disrupt expression of a cellular KIM gene,
 CC also useful in therapy (e.g. to block the growth of tumours dependent on
 CC KIM for growth) or compositions. The proteins and polynucleotides are
 CC useful diagnostically e.g. to detect and quantify renal injury/disease
 CC (indicative of increased risk, or presence of, renal injury or impaired
 CC function), or abnormal responses to tissue injury (indicative of
 CC renal dysfunction).
 CC tissue growth arising from/affecting renal tissue) The proteins can
 CC also be used to locate KIM-producing cells (especially specific loci,
 CC e.g. tissue masses abnormally producing/expressing KIM such as tumours
 CC arising from/affecting renal tissue), by contacting cells with an
 CC imageable KIM-binding reagent and imaging reagent accumulation.
 XX Sequence 339 AA;
 QY Query Match 94.8%; Score 182; DB 20; Length 339;
 Beat Local Similarity 91.2%; Pred. No. 1.2e-16;
 Matches 31; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 DB 185 RRSRYKLLNMAVYOOVQONKEDAMTEHDVWMEI 218
 QY 1 RRSRYKLLNMAVYOOVQONKEDAMTEHDVWMEI 34
 ID AAM79829 standard; Protein; 374 AA.
 XX AAM79829;
 AC AAM79829;
 XX

DT 06-NOV-2001 (first entry)
 XX Human protein SEQ ID NO 3475.
 XX Human, cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX Homo sapiens.
 XX MO200157199-A2.
 XX 09-AUG-2001.
 PD 05-FEB-2001; 2001MO-US04098.
 XX 03-FEB-2000; 2000US-046614.
 XX 27-APR-2000; 2000US-0560875.
 XX 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0520325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX (HYSB-) HYSB INC.
 XX Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y,
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 XI Xie AJ, Yang Y, Wehrman T, Goodrich R;
 XX NPI; 2001-476283/51.
 DR N-PSDB; AAK52582.
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX Claim 20; Page 350; 6223pp; English.
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC proteins encoded by them (AAK51456-AAK53435) that exhibit activity relating to
 CC cytokine, cell proliferation or other cell population. They may induce
 CC production of other cytokines in other cell populations. The KIM
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity in inhibiting activity, and may be used in the treatment of
 CC inflammation, cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAK60020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX Sequence 374 AA;
 QY Query Match 93.2%; Score 179; DB 22; Length 374;
 Beat Local Similarity 97.1%; Pred. No. 3.3e-16;
 Matches 33; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 DB 220 RRSRYKLLNMAVYOOVQONKEDAMTEHDVWMEI 253
 QY 1 RRSRYKLLNMAVYOOVQONKEDAMTEHDVWMEI 34
 ID AAB49483 standard; Protein; 488 AA.
 XX AAB49483;
 AC AAB49483;
 XX 08-MAR-2001 (first entry)
 DT

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 16:43:45 / Search time 23.0233 Seconds

(without alignments) 196,889 Million cell updates/sec

Title: US-10-010-667a-19

Sequence: 1 REVHPLATSHOQFYKPIILV 22

Scoring table: BL0SUM62

Gapop 10.0, Gapext 0.5

Search: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum March 08

Database: SPTRMBL_21.*

1: sp_Archaea:*
2: sp_Bacteria:*
3: sp_Fungi:*
4: sp_Human:*
5: sp_Invertebrate:*
6: sp_Mammal:*
7: sp_Thc:*
8: sp_Organelle:*
9: sp_Phage:*
10: sp_Plant:*
11: sp_Rodent:*
12: sp_Virus:*
13: sp_Unclassified:*
14: sp_Yeast:*
15: sp_Yrivers:*
16: sp_Bacteriap:*
17: sp_Archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	116	100.0	264	4	Q9Y6U5	Q9Y6U5 homo sapien
2	116	100.0	264	4	Q9Y6U5	Q9Y6U5 homo sapien
3	94	81.0	339	11	Q9CMR7	Q9CMR7 mus musculu
4	94	81.0	339	11	Q924J2	Q924J2 mus musculu
5	94	81.0	339	11	Q924J9	Q924J9 mus musculu
6	48	41.4	172	2	Q93BL7	Q93BL7 streptococc
7	48	41.4	388	16	Q9A0S6	Q9A0S6 streptococc
8	48	41.4	456	4	Q8TDP3	Q8TDP3 homo sapien
9	48	41.4	456	4	Q8TDP3	Q8TDP3 homo sapien
10	48	41.4	456	4	Q8TDP3	Q8TDP3 homo sapien
11	48	41.4	456	4	Q8TDP3	Q8TDP3 homo sapien
12	47	40.5	233	10	Q90AP4	Q90AP4 carid hept
13	47	40.5	427	16	Q8RH72	Q8RH72 themanact
14	47	40.5	431	17	Q9HMR3	Q9HMR3 halobacteri
15	47	40.5	469	4	Q9C0P0	Q9C0P0 homo sapien
16	47	40.5	565	3	Q04600	Q04600 methanocarc

17	47	40.5	568	4	Q96M03	Q96M03 homo sapien
18	47	40.5	573	3	Q90M38	Q90M38 schizosacch
19	47	40.5	573	3	Q90M38	Q90M38 schizosacch
20	47	40.5	763	10	Q91M12	Q91M12 oryza sativ
21	47	40.5	763	10	Q91M12	Q91M12 oryza sativ
22	47	40.5	806	10	Q91M22	Q91M22 oryza sativ
23	47	40.5	912	12	Q80V01	Q80V01 buffalo her
24	47	40.5	947	12	Q90AP7	Q90AP7 bovine hept
25	46	40.1	437	2	Q91YV7	Q91YV7 glycine max
26	46	39.7	169	10	Q91B65	Q91B65 glycine max
27	46	39.7	296	16	Q8YH82	Q8YH82 listeria mo
28	46	39.7	296	16	Q8YH82	Q8YH82 listeria mo
29	46	39.7	425	16	Q92E57	Q92E57 listeria in
30	46	39.7	425	16	Q92E57	Q92E57 listeria in
31	46	39.7	430	16	Q93P52	Q93P52 bacillus ha
32	46	39.7	473	11	Q930J6	Q930J6 mus musculu
33	46	39.7	488	11	Q93P41	Q93P41 ratuora norv
34	46	39.7	509	11	Q93X25	Q93X25 mus musculu
35	46	39.7	735	16	Q82D01	Q82D01 versatilis pe
36	46	39.7	735	16	Q82D01	Q82D01 versatilis pe
37	46	39.7	931	12	Q90AP6	Q90AP6 rangiferine
38	45	38.8	151	5	Q9M5A2	Q9M5A2 dirosophila
39	45	38.8	151	5	Q9M5A2	Q9M5A2 dirosophila
40	45	38.8	428	10	Q9W2P5	Q9W2P5 arabidopsis
41	45	38.8	436	10	Q9W2P5	Q9W2P5 arabidopsis
42	45	38.8	436	10	Q9W2P5	Q9W2P5 arabidopsis
43	45	38.8	436	10	Q9W2P5	Q9W2P5 arabidopsis
44	45	38.8	919	12	Q90AP5	Q90AP5 caprine her
45	44.5	38.4	308	17	Q87H12	Q87H12 methanocarc

ALIGNMENTS

RESULT 1

Q9Y6U5 PRELIMINARY; PRT; 264 AA.

AC Q9Y6U5; 01-NOV-1999 (TRENBERL. 12, Created)
DT 01-NOV-1999 (TRENBERL. 12, Last sequence update)
DT 01-DEC-2001 (TRENBERL. 19, Last annotation update)
DE Q9Y6U5 Homo sapiens (human)
OS Homo sapiens (human)
OC Eukaryota; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE:9363792; PubMed:947074;
RT "Homo sapiens BAC clone CTR-8/515."
RL Genome Ref. 8:1097-1108(1998).
[2]
RN SEQUENCE FROM N.A.
RT Strong C., Layman D., Stowman C.,
RA "The sequence of Homo sapiens BAC clone CTR-8/515."
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; Acc05061; AA043182.1; -.
FT NON_TER
FT 1
FT 264
SQ SEQUENCE 264 AA, 31203 MW, 35C948300357872 CRC64;

Query Match 100.0%; Score 116; DB 4; Length 264;

Best Local Similarity 100.0%; Pred. No. 3.1e-11; Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHOQFYKPIILV 22


```

RA MEDLINE-21992816; PubMed-21997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Chae Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA RT "A complete sequence of T. tengcongensis genome."
RA Genome Res. 12:689-700(2002).
RL EMBL; AB033162; AAM2314.1; -.
KV Complete proteome.
SQ SEQUENCE 427 AA; 46642 MW; 0AAEF796CE84B CRC64;

Query Match 40.5%; Score 47; DB 16; Length 427;
Best Local Similarity 72.7%; Pred. No. 17;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VHPPLATSHQ 13
Db 376 VHPPLATSHQ 386

RESULT 14
ID 09HMR3 PRELIMINARY; PRT; 431 AA.
AC 09HMR3;
DT 01-MAR-2001 (TRENBERG, 16, Last sequence update)
DT 01-MAR-2002 (TRENBERG, 16, Last annotation update)
DS 09HMR3;
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN 11)
RP SEQUENCE FROM N.A.
RX MEDLINE-250463; PubMed-11015950;
RA M. W. Krumholz, S. P. Mahapatra, C. G. Bertoliet, B. Pan M.,
RA Shukla H. D., Lashy S. R., Baliga N. S., Thomson V., Sirogna J.,
RA Leitbauer B., Keller K., Cruz R., Danson M. J., Hough D. W.,
RA Maddocks D. G., Jablonski P. E., Krebs M. P., Angevine C. M., Dale H.,
RA Jendryaszek T. A., Peck R. P., Pohlschroder M., Spulich J. L., Jung K.-H.,
RA Alan M., Freitas T., Hou S., Daniels C. J., Dennis P. P., Omer A. D.,
RA Zharhik H., Lowe T. N., Liang F., Riley K., Hood L., Dasarma S.;
RA RT "Complete genome sequence of the extreme halophile, Halobacterium
RA Proc Natl Acad Sci U S A. 97:12176-12181(2000).
DR EMBL; AB005123; AAC20508.1; -.
DR HESP; P00935; ICS1.
DR InterPro; IP000277; Cys_Met_Meta_PP.
DR Pfam; PF01053; Cys_Met_Meta_PP; 1.
KV Complete proteome.
SQ SEQUENCE 431 AA; 45768 MW; 67CEC5064A2DE9B3 CRC64;

Query Match 40.5%; Score 47; DB 17; Length 431;
Best Local Similarity 72.7%; Pred. No. 17;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VHPPLATSHQ 13
Db 376 VHPPLATSHQ 386

RESULT 15
ID 09C0F0 PRELIMINARY; PRT; 469 AA.
AC 09C0F0;
DT 01-JUN-2001 (TRENBERG, 17, Created)
DT 01-JUN-2001 (TRENBERG, 17, Last sequence update)
DT 01-OCT-2001 (TRENBERG, 18, Last annotation update)
DE KIAA1713 protein (Fragment).
GN KIAA1713.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

```

```

OX NCBI_TaxID=9606;
RN 11)
RP SEQUENCE FROM N.A.
RX MEDLINE-21082932; PubMed-11214970;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RA RT "Prediction of the coding sequences of unidentified human genes. XIX.
RA The complete sequences of 100 new cDNA clones from brain which code
RA for large proteins in vitro."
RL DNA Res. 7:347-355(2000).
RT DNA Res. 7:347-355(2000).
DR EMBL; AB051500; BAB21804.1; -.
DR InterPro; IP003380; H3_Proteasome_rec.
DR HESP; P01471; HISTATIN3R.
DR PRINTS; PR01471; HISTATIN3R.
DR NON TER 1
SQ SEQUENCE 469 AA; 50649 MW; 235F3AF4E627525 CRC64;

Query Match 40.5%; Score 47; DB 4; Length 469;
Best Local Similarity 59.3%; Pred. No. 13;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 9 TSHQYFYKIP 20
Db 146 TSHQYFYKIPV 157

Search completed: March 26, 2003, 16:49:44
Job time : 25.0233 secs

```




```

Query Match Summary      42.2% Score 49; DB 2; Length 177;
Best Local Similarity    50.0% Pred. No. 2.1/
Matches      9; Conservative      1; Mismatches      8; Gaps      0;
Oy      2 EYHPLATSHNOOVPKYP 19
      | : | | | | | | | | |
Db      135 EYHPLATSHNOOVPKYP 152

RESULT 3
C:cytochrome f - red alga (Porphyra purpurea) chloroplast
C:species: chloroplast Porphyra purpurea
C:date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 11-Jun-1999
C:accession: S73186
R:reith, M.; Muniholland, J.
Plant. Mol. Biol. Rep. 13; 333-335, 1995
A:accession: S73186
A:definition: The sequence of the Porphyra purpurea chloroplast genome.
A:status: preliminary
A:submitter: preliminary
A:sequence_type: DNA
A:accession: 1-1330 -REI-
A:cross-references: EMBL:U08804; NID:G1276653; PTDN:AA00151.1; PID:G1276731
A:comment: The nucleotide sequence was submitted to the EMBL data library, October 1995
A:clone: chloroplast
A:genome: chloroplast
C:superfamily: cytochrome f
C:keywords: chloroplast; heme; thylakoid

```

```

Query Match      40.9%   Ped. 40.7 %; Db 2; Length 320;
Best Local Similarity 40.7%   Fed. No. 7.1;
Matches 11; Conservative 4; Mismatches 7; Indels 5; Gaps 1;

Cy          1 RUTPLR-----ATSHOQKPKPDI 22
||| || | ||| ||| ||| ||| ||| ||| |||
Db         160 RLTFPLSPDAKDKQHFFRVIVT 166

RESULT 4
De4797
Hypothetical protein Atc2g37810 [Imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
Accession: De4797 #sequence_revision 02-Feb-2001 #ext_change 02-Feb-2001
Citation: Rilling, X.; Kaul, S.; Ronaldson, S.D.; Shea, T.P.; Bantito, M.I.; Town, C.D.; Fujii, C.Y.;
Liu, H.; Koob, H.; Moffett, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.L.; Tallon, L.;
Gusis, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, A.
Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: Ab4420; NCBI:20083487; PMID:10617197
A/Accession: DB4797
A/Accession: PEP4797
A/Molecule type: Protein
A/Residue: 1+233 ->STO-
A/Cross-references: GB:AE002093; NID:s4895203; PINR:AA032790.1; GSDB:GN00139
A/Gene: Atc2g37810
C/Map position: 2
```

Query Match	40.5%	Score 67	DB 2	Length 233
Best Local Similarity	43.8%	Pred. No. 6		
Matches	7	Conservative	3	Mismatches
			Indels	0
			Gaps	0
QY	4	IHPATSHOOTFKAP	19	
	:	: :: :		
b5	105	VHPLCTOHQHVSVP	120	

RESULT 5
H84392

Q:acetyl_homoserine [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
G:Accession: H64352
R:HighThroughput: S.P., Mahaffey, G.G.; Benayahu, B.; Pan, M.; Shukla, H.D.; Lasky, S.
R:HighThroughput: Miller, K.; Cruz, T. R.; Danson, M.J.; Hough, D.W.; Madocke, D.G.; Jablon
R:Jung, K.H.; Alam, M.; Fretz, T.
P:Proc. Natl. Acad. Sci. U.S.A. 97, 12376-12381, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Emer, A.D.; Ehrhardt, H.; Lowe, M.; Little
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160, MIMD:20504483, PMID:11016950
A:Accession: H64352
A:Accession: H64352
A:Molecule: cDNA
A:Molecule: cDNA
A:Protein-references: 15431 ->O
A:Cross-references: GB:AE004437; NID:G10581828; PIRN:AA020508.1; GSPDB:GN00136
C:Gene: hsl
C:Gene: hsl
C:Superfamily: O-succinylhomoserine (thiol)-lyase
Query Match 40.5%, Score 47, DB 2, Length 431;
Best Local Similarity 72.7%, Pred. NO. 12;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

RESULT 6
SS2682
Hypochemical protein YDR117c - Yeast (*Saccharomyces cerevisiae*)
Nucleotide names: Hypochemical protein ID9727.12c-
C.date: 19-May-1995 #sequence_revision 01-Sep-1995 #text_change 19-Apr-2002
C.accession: SS2682
R.Murphy, L., Shore, L., Harris, D.
submitted to the EMBL Data Library, March 1995
A.Reference number: SS2671
A.accession: SS2682
A.molecule type: DNA
A.sequence length: 1587
A.cross-references: EMBL:Z48759; NID:G747879; PID:G747891; GSPDB:GSM00004; MIPS:YDR117C
C.gene: MIPS:YDR117C
A.gene: MIPS:YDR117C
A.cross-references: SCD:S0002524
Map position: 48

```
OY      2 EYHPLATSHOQTYTXI 18  
          ||::|||:::|6:  
Db       439 EYLHPLATNNFTETRYQL 455  
  
Query Match      40.5% Score 47; DB 2; Length 565;  
Best Local Similarity 47.1%; Pred. No. 16;  
Matches    8; Conservative   5; Mismatches    4; Indels    0; Gaps
```

RESULT 7

T3J9727
nucleosporin homolog - fission yeast (*Schizosaccharomyces pombe*)
[C:\SpecGen\Blast\blastall.bat](#) poms1.tbl -fission01-dec-1999 %text_change 07-dec-1999
T3J9727 %sequence %revision 01-dec-1999
Claccense: T3J9727
R:Mood, V.; Rajandream, M.A.; Barrell, B.G.; Volkart, G.
submitted to the EMBL Data Library, August 1999
%reference number: Z21874

A:status: preliminary; translated from GB/EMBL/DBJ
A:molecule type: DNA
A:residues: 1-675 <MO>
A:cross-reference: EMBL:AL10946; PIDN:CBAS2802.1; GSPDB:GN00067; SPDB:SPCK709.04c
A:experimental source: EMBL:strain 972h; cosmid c1759
A:genetics:

A:Gene: SPDB:SPEC17G9.04c
A:Map position: 2

Query Match
Best Local Similarity 40.5%; Score 47; DB 2; Length 675;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 VHPPLATSHOQYFV 16
DB 566 IISPLATSHOQYFV 579

RESULT 8
T07623
extensin homolog HRGP2 - soybean (fragment)
N:Alternate names: hydroxyproline-rich glycoprotein HRGP2
C:Species: Glycine max (soybean)
C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 11-May-2000
C:Accession: T07623
A:Authors: Kretz, J.; Kim, M.; Kunst, F.; Kurupkac, G.; Madueno, E.; Maitounam, A.; Mat
ok, C.; Schluecker, T.; Simoes, N.; Tieretz, A.; Vazquez-Boland, J.A.; Voess, H.; Wehland,
A. Title: Isolation and characterization of three soybean extensin cDNAs.
A:Reference number: Z16058; PMID:94211912; PMID:8159793
A:Accession: T07623
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-169 -CHMU-1122010, NID:G347454, PIDN:AB033971.1; PID:G347455
A:Experimental source: strain Wayne; seedling
C:Genetics:
A:Gene: HRGP2
C:Superfamily: hydroxyproline-rich glycoprotein
C:Keywords: glycoprotein; hydroxyproline

Query Match
Best Local Similarity 39.7%; Score 46; DB 2; Length 169;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 6 PLATSHOQYFV 19
DB 92 PPSPLSHOQYFV 105

RESULT 9
A11178
transcription regulator homolog lmo083 [imported] - *Listeria monocytogenes* (strain EGD-
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: A11178
A:Authors: Kretz, J.; Kim, M.; Kunst, F.; Kurupkac, G.; Madueno, E.; Maitounam, A.; Mat
ok, C.; Schluecker, T.; Simoes, N.; Tieretz, A.; Vazquez-Boland, J.A.; Voess, H.; Wehland,
A. Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; PMID:21537279; PMID:11679669
A:Accession: A11178
A:Status: preliminary
A:Residues: 1-236 -CHMU-1122010, NID:G347454, PIDN:AB033971.1; PID:G347455
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo083

Query Match
Best Local Similarity 39.7%; Score 46; DB 2; Length 236;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 2 EVHPLATSHOQYFV 18
DB 222 IISPLATSHOQYFV 238

RESULT 10

AD1508
O-acetylhomoserine sulphydrolase homolog lmo0595 [imported] - *Listeria monocytogenes* (str
C:Species: *Listeria monocytogenes*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AD1508
A:Authors: Kretz, J.; Kim, M.; Kunst, F.; Kurupkac, G.; Madueno, E.; Maitounam, A.; Mat
ok, C.; Schluecker, T.; Simoes, N.; Tieretz, A.; Vazquez-Boland, J.A.; Voess, H.; Wehland,
A. Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; PMID:21537279; PMID:11679669
A:Accession: AD1508
A:Status: preliminary
A:Residues: 1-425 -CHMU-1122010, NID:G347454, PIDN:AB033971.1; PID:G347455
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo0595
C:Superfamily: O-succinylhomoserine (thiol)-lyase

Query Match
Best Local Similarity 39.7%; Score 46; DB 2; Length 425;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VHPPLATSHOQ 13
DB 376 IISPLATSHOQ 386

RESULT 11
AD1508
O-acetylhomoserine sulphydrolase homolog lmo064 [imported] - *Listeria innocua* (strain C)
C:Species: *Listeria innocua*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AD1508
A:Authors: Kretz, J.; Kim, M.; Kunst, F.; Kurupkac, G.; Madueno, E.; Maitounam, A.; Mat
ok, C.; Schluecker, T.; Simoes, N.; Tieretz, A.; Vazquez-Boland, J.A.; Voess, H.; Wehland,
A. Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; PMID:21537279; PMID:11679669
A:Accession: AD1508
A:Status: preliminary
A:Residues: 1-425 -CHMU-1122010, NID:G347454, PIDN:AB033971.1; PID:G347455
A:Experimental source: strain C11p1262
C:Genetics:
A:Gene: lmo064
C:Superfamily: O-succinylhomoserine (thiol)-lyase

Query Match
Best Local Similarity 39.7%; Score 46; DB 2; Length 425;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VHPPLATSHOQ 13
DB 376 IISPLATSHOQ 386

RESULT 12
T06576
phenylalanine-tRNA ligase (EC 6.1.1.20) beta chain T22F9.180 [similarity] - *Arabidopsis t*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Jun-2002

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 16:38:14 ; Search time 42.7209 Seconds

(without alignments) 68.620 Million call updates/sec

Title: US-10-010-667a-19

Sequence: 1 REVHPLATSHQVFXKPIILV 22

Scoring table: BLOSUM62

Gapco 10.0 , Gapext 0.5

Searched: 908470 segs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum March 08

Maximum March 1008

Listing first 45 summaries

Database : A_GeneSeq 101002:*

1: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*

2: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*

3: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*

4: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*

5: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*

6: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*

7: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*

8: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*

9: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*

10: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*

11: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*

12: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*

13: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*

14: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*

15: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*

16: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*

17: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*

18: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*

19: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*

20: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*

21: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*

22: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*

23: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score distribution, and is derived by analysis of the score score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	22	21	AAV58198 Human STRAP-1 pept
2	116	100.0	22	22	AAE02786 Extracellular loop
3	116	100.0	104	22	AAE02786 Extracellular loop
4	116	100.0	104	22	AAE02786 Extracellular loop
5	116	100.0	104	21	ABG43822 Human peptide enco
6	116	100.0	339	21	AAV58198 Human STRAP-1 proc
7	116	100.0	339	22	AAU65927 Human prostate CDN
8	116	100.0	339	22	AAU78845 Human prostate CDN
9	116	100.0	339	22	AAU01282 P789p amino acid 8
10	116	100.0	339	23	ABG61813 Prostate cancer-ab

11	116	100.0	339	23	ABB95387 Human P789p protei
12	116	100.0	374	22	AAW79829 Human protein SEQ
13	116	100.0	375	22	AAE02786 Human S1X-Transmem
14	116	100.0	375	22	AAE02786 Human S1X-Transmem
15	89	79.3	109	20	AAU12304 Human S1X-Transmem
16	89	76.7	339	20	AAE02789 Kidney Injury asso
17	68	58.6	26	22	AAE02789 Extracellular loop
18	68	58.6	419	22	AAU010189 Human ORF3 of S1X
19	68	58.6	450	22	AAE02841 Human STRAP-2 proc
20	68	58.6	454	22	AAU010188 Human ORF2 of S1X
21	68	58.6	454	22	AAE01591 Human S1X-Transmem
22	68	58.6	454	22	AAE01591 Human S1X-Transmem
23	68	58.6	454	23	AAU080130 Human PUMPCH prote
24	68	58.6	454	23	AAU080130 Human PUMPCH prote
25	68	58.6	490	22	AAU010187 Human S1X-Transmem
26	68	58.6	576	22	ABG12306 Novel human diagno
27	68	58.6	1273	22	AAU060113 Human polyepitide
28	49	42.2	129	22	AAU010072 Human polyepitide
29	48	41.4	328	22	ABP23255 Human polyepitide
30	48	41.4	456	22	ABP23255 Human polyepitide
31	48	41.4	456	22	ABP23255 Human polyepitide
32	48	41.4	487	22	ABP23255 Human polyepitide
33	48	41.4	488	22	AAU10220 Human P-HYDE. Hom
34	48	41.4	488	22	AAU04564 Human S1X-Transmem
35	48	41.4	488	22	AAU04564 Human S1X-Transmem
36	48	41.4	488	22	AAU04564 Human S1X-Transmem
37	48	41.4	488	22	AAU04564 Human S1X-Transmem
38	48	41.4	488	22	AAU04564 Human S1X-Transmem
39	46	39.7	296	23	ABP47546 Human protein sequ
40	46	39.7	425	23	ABP47546 Human protein sequ
41	46	39.7	488	22	AAU04564 Human S1X-Transmem
42	46	39.7	526	23	ABP47546 Human protein sequ
43	45	38.8	146	22	ABP47546 Human protein sequ
44	45	38.8	146	22	ABP47546 Human protein sequ
45	44.5	38.4	1176	22	ABP47546 Human protein sequ

ALIGNMENTS

RESULT 1	AAV58198	AAV58198 standard; peptide, 22 AA.
AC	AAV58198;	
XX		
DT	14-MAR-2000 (first entry)	
DE		
XX	Human STRAP-1 peptide, corresponding to STRAP-1 extracellular region 1.	
KW	Scraping transmembrane antigen of the prostate; STRAP-1; prostate;	
KW	transmembrane domain; type IIIa membrane protein; expression; cancer;	
KW	prostate cancer; bladder cancer; colon cancer; pancreatic cancer;	
KW	cellular; humoral; anticancer vaccine; antibody; detection; diagnosis;	
KW	prognosis; monitoring; susceptibility; therapeutic inhibitor;	
KW	drug targeting; recombinant protein.	
OS	Synthetic.	
XX		
XX	Homologous.	
PN	MO962941-A2.	
XX		
PD	09-DEC-1999.	
XX		
XX	01-JUN-1999, 99MO-US12157.	
PR	01-JUN-1998, 98US-0087520.	
PR	30-JUN-1998, 98US-0081183.	
XX		
XX	(UROG-) UROGENESIS INC.	
PA	(AFAR) AFAR D E.	
PA	(HUBB) HUBB R S.	

PA	(LEON ¹) LEONG K.
PA	(BAIT ²) RAITANO A B.
PA	(SAFP ³) SAEFMAN D C.
PI	
PI	Akar DE, Hubert RS, Leong K, Raitano AB, Saffran DC,
DR	WPI, 2000-07832/06.
XX	
PT	Novel proteins useful as diagnostic markers and therapeutic targets,
PT	particularly for prostate cancer -
PX	
XX	
XX	Disclosure: Page 22; 83pp. English.
CC	Sequence AY51918-Y5620 represent synthetic peptides that correspond
CC	to the extracellular regions of STRAP-1 (serpinin transmembrane
CC	antigen of the prostate, AA51914). These peptides were used to raise
CC	monoclonal anti-STRAP-1 antibodies. STRAP-1 is the prototype
CC	member of the STRAP family of proteins (AA51914-Y5819) which
CC	exhibit a high degree of structural conservation, but which show
CC	no significant homology to other serpins. The expression of STRAP-1
CC	has not been localized to chromosome 7p22. STRAP-1 is thought to be a
CC	type IIa membrane protein and is expressed predominantly in prostate
CC	cells in normal human tissues. Structurally, STRAP-1 is a 339 amino
CC	acid protein characterized by six transmembrane domains and
CC	intracellular N- and C-termiin, suggesting that it folds in a
CC	"serpinin" manner into three extracellular and two intracellular loops.
CC	STRAP-1 mRNA and protein expression is maintained at high levels and is
CC	also overexpressed in certain other cancers including bladder,colon, is
CC	pancreatic and ovarian cancer. The function of the STRAP proteins is not
CC	known. They may be ion channels (from the presence of six transmembrane
CC	domains), a feature which is shared by certain ion channels) or
CC	gap-junction proteins (from immunohistochemical staining). STRAP-1 and
CC	STRAP-2 are cell-surface tumour antigens. Immunisation with a STRAP
CC	peptide containing cellular STRAP immunotumour associated antigenically
CC	specific-binding sites, to produce anticancer vaccines and to generate
CC	specific binding agents. The antibodies may be used for detection, prognosis,
CC	and monitoring of cancers (or susceptibility to cancer), as therapeutic,
CC	inhibitors or to target therapeutic agents to their site of action. STRAP
CC	nucleotide acids may be used for recombinant protein production, as
CC	diagnostic and prognostic reagents, for identifying STRAP-expressing
CC	cell-lines, and for studying the role of STRAP expression in disease. High levels of STRAP
CC	proteins are exposed on the cell surface. They are easily targeted by
CC	systemically administered agents, and because they are expressed mainly
CC	on prostatic epithelial cells, agents targetted to them should have
CC	minimal side effects on other tissues.
XX	
XX	Sequence 22 AA:
XX	
XX	Query Match 100.0%; Score 116; DB 21; Length 22;
XX	Best Local Similarity 100.0%; Pctd. No. 7e+12;
XX	Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
DY	1 REVHPLATSHOOTFKPKPILV 22
DY	. 1 REVHPLHTSHOOFKPKPILV 22
XX	
RESULT 2	
AAE02766	
ID	AAE02766 standard; peptide; 22 AA.
AC	AAE02766;
DC	06-AUG-2001 (first entry)
DE	Extracellular Loop #1 of human STRAP-1, suitable for cloning into pC-
XX	
XX	Human; cytosolic; antiproliferative; vaccine; gene therapy;
XX	six transmembrane epithelial antigen of the prostate-1; STRAP-1;
XX	chromosome 7p22.3; cancer; prostate; colon;bladder; pancreatic; lung;

XX	ovarian, extracellular loop; serpentine transmembrane antigen.
KW	Homo sapiens.
XX	
PN	MO200140276-A2.
PD	07-JUN-2001.
XX	
PF	06-DEC-2000; 2000MO-US33040.
XX	
XX	06-DEC-1999; 99US-0455486.
XX	
PA	(URGC-) UROGENESIS INC.
PI	Afar EH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC, Faris M, Jakobovits A.
XX	
DR	WPI, 2001-367804/38.
XX	
XX	New STEAP (six transmembrane epithelial antigen of the prostate) protein, expressed in human cancers, useful for detecting and treating cancer -
PT	
PT	
XX	
XX	Example 19, Page 102, 187p; English.
XX	
CC	The present invention relates to human six transmembrane epithelial antigen of the prostate (STEAP) protein, a member of the STEAP-1 gene family. The present invention provides a STEAP-1 gene
CC	therapy, inhibiting the development or progression of a cancer (e.g. prostate, colon, bladder, lung, ovarian and pancreatic) expressing STEAP
CC	or inhibiting growth or killing cells expressing STEAP in a patient, treating a patient with a cancer that expresses STEAP, or inhibiting growth or
CC	killing cells expressing STEAP, comprises administering to the patient a
CC	variable domain of the heavy and light chains of the monoclonal antibody that specifically binds to STEAP, such that the vector delivers the
CC	single chain monoclonal antibody coding sequence to the cancer cells and
CC	the encoded single chain monoclonal antibody is expressed
CC	intracellularly. The present sequence is extracellular loop of STEAP-1
CC	suitable for cloning into pFC, which is used in the invention. STEAP-1
XX	gene is located on chromosome P22.3.
XX	
XX	Sequence 22 AA:
XX	
Query Match	100.0%; Score 116; DB 22; Length 22;
Beat Local Similarity	100.0%; Pred. No. 76-12;
Matches 22; Conservative 0; Mismatches	0; Indels 0; Gaps 0;
Y	1 REVHPLATSHQVFKFKPIIV 22
DB	1 REVHPLATSHQVFKFKPIIV 22
XX	
XX	RESULT 3
ID	ABB40417
ID	ABB40417 standard; Peptide; 104 AA.
XX	
XX	ABB40417,
DT	04-FEB-2002 (first entry)
XX	
DE	Peptide #1923 encoded by human foetal liver single exon probe.
KW	Human; foetal liver; gene expression; single exon nucleic acid probe.
XX	
XX	Homo sapiens.
XX	
PN	MO200157277-A2.
XX	
XX	09-AUG-2001.
XX	
XX	30-JAN-2001; 2001MO-US00669.


```

XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 03-AUG-2000; 2000US-0232366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
XX
XX Claim 27; SEQ ID NO 33053; 639bp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX fetal liver. The present sequence is a peptide encoded by a single exon
XX nucleic acid probe of the invention.
XX
XX Note: The sequence data for this patent did not form part of the
XX sequence data submitted to the EMBL database. The format directly
XX from WIGO at ftp.wigo.int/pub/published_pat_sequences.
XX
XX Sequence 104 AA;
XX
XX Query Match 100.0%; Score 116; DB 22; Length 104;
XX Best Local Similarity 100.0%; Pred. No. 3, 8e-11;
XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 REVHPLATSHQQYFKKPIPLV 22
XX |||||||||||||||||||
XX 59 REVHPLATSHQQYFKKPIPLV 80
XX
XX RESULT 4
XX ID AAM73944
XX AAAM73944 standard; Protein; 104 AA.
XX
XX AAM73944;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 34250.
XX
XX Human bone marrow expressed exon, gene expression analysis; probe;
XX microarray; cancer; leukemia; lymphoma; myeloma.
XX
XX Homo sapiens.
XX
XX NO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001MO-US00668.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 01-AUG-2000; 2000US-0632366.
XX 02-AUG-2000; 2000US-0634866.
XX 27-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX

```

```

XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 34250; 658bp + sequence listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to predict, measure and display gene
XX expression in samples derived from human bone marrow. The present
XX sequence is a peptide encoded by one of the probes of the invention.
XX
XX Sequence 104 AA;
XX
XX Query Match 100.0%; Score 116; DB 22; Length 104;
XX Best Local Similarity 100.0%; Pred. No. 3, 8e-11;
XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 REVHPLATSHQQYFKKPIPLV 22
XX |||||||||||||||||||
XX 59 REVHPLATSHQQYFKKPIPLV 80
XX
XX RESULT 5
XX ID ABG43832
XX ABG43832 standard; Peptide; 104 AA.
XX
XX ABG43832;
XX
XX 19-AUG-2002 (first entry)
XX
XX Human peptide encoded by genome-derived single exon probe SEQ ID 33497.
XX
XX Human, single exon probe; asthma; lung cancer; COPD; IHD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hereditary Rüdik syndrome; sarcoidosis; pulmonary haemorrhoidosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
XX primary biliary cirrhosis; primary biliary dyskinesia;
XX hyaline membrane disease.
XX
XX Homo sapiens.
XX
XX NO200186003-A2.
XX
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001MO-US00665.
XX
XX 04-FEB-2000; 2000US-180312P.
XX 26-MAY-2000; 2000US-207456P.
XX 30-JUN-2000; 2000US-0608408.
XX 01-AUG-2000; 2000US-0632366.
XX 02-AUG-2000; 2000US-0634866.
XX 27-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples -
XX
XX Claim 27; SEQ ID NO 33497; 634bp; English.
XX

```


CC STRAP-expressing cells. STRAP proteins may be used to identify
 CC specific-binding agents, to produce anticancer vaccines and to generate
 CC specific antibodies. The antibodies may be used for detection, prognosis,
 CC therapy and/or prevention of cancer. The antibodies may be used to
 CC inhibit or to target therapeutic agents to their site of action. STRAP
 CC inhibitors or to target therapeutic agents to their site of action. STRAP
 CC diagnostic acid may be used for recombinant protein production, as
 CC cells for screening inhibitors of STRAP expression and for therapeutic
 CC modulation/inhibition of STRAP expression. Since high levels of STRAP
 CC proteins are exposed on the cell surface, they are easily targeted by
 CC systemically administered agents, and because they are expressed mainly
 CC on the cell surface, they are easily targeted to them should have
 CC animal side effects on other tissues.

XX Sequence 339 AA;
 SQ
 Query Match 100.0%; Score 116; DB 21; Length 339;
 Best Local Similarity 100.0%; Pred. No. 1,4e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 REVHPLATSHOQYFKPIPLV 22
 DB 92 REVHPLATSHOQYFKPIPLV 113

RESULT 7
 AA069927
 AA069927 standard; Protein, 339 AA.
 AA069927;
 30-JAN-2002 (first entry)
 Human prostate cDNA encoded protein #72.
 Human prostate cancer, cytostatic; immunostimulant; tumour; immunogen.
 Homo sapiens.
 NO200173032-A2.
 PN
 PD 04-OCT-2001.
 XX
 XX 27-MAR-2001; 2001NO-US09319.
 XX
 XX 27-MAR-2001; 2000US-0536857.
 PR 09-MAY-2000; 2000US-0568100.
 PR 12-MAY-2000; 2000US-0570737.
 PR 13-JUN-2000; 2000US-0593793.
 PR 27-JUN-2000; 2000US-0605783.
 PR 18-AUG-2000; 2000US-0638215.
 PR 18-AUG-2000; 2000US-0638215.
 PR 06-SEP-2000; 2000US-0652719.
 PR 02-OCT-2000; 2000US-0679466.
 PR 10-OCT-2000; 2000US-0685166.
 XX
 PA (CORI-) CORIXA CORP.
 PI
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kaloe MD;
 PI W. J. Harlocker MD, J. K. Harlocker MD, J. K. Harlocker MD, J. K.
 PI Li SX, Wang A, Skeiky VM, Hopfer WT, Henderson RA,
 DR N-PSDB; AAS64160.
 XX
 XX New human prostate-specific polypeptides and polynucleotides useful for
 XX the diagnosis and treatment of cancer, especially prostate cancer -
 XX Claim 20; Page 549; 579p; English.

CC The invention relates to isolated prostate-specific
 CC polynucleotides, polypeptides, fusion proteins of the polypeptides,
 CC antibodies raised against the polypeptides (or antigenic epitopes

CC derived from them) and antigen-presenting cells expressing the
 CC polypeptides. The antibodies are useful for detecting the presence of
 CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
 CC fusion proteins are useful for the diagnosis and treatment of cancer.
 CC T cells specific for tumor protein, and for inhibiting the development
 CC of cancer especially prostate cancer. Compositions comprising the
 CC polynucleotide and/or polypeptide are useful for stimulating an immune
 CC response, and for treating cancer. The oligonucleotide is useful for
 CC detecting cancer. The present sequence is a prostate specific
 CC polypeptide of the invention.

XX Sequence 339 AA;
 SQ
 Query Match 100.0%; Score 116; DB 22; Length 339;
 Best Local Similarity 100.0%; Pred. No. 1,4e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 REVHPLATSHOQYFKPIPLV 22
 DB 92 REVHPLATSHOQYFKPIPLV 113

RESULT 8
 AA078845
 AA078845 standard; Protein, 339 AA.
 AA078845;
 06-NOV-2001 (first entry)
 Human protein SEQ ID NO 1507.
 Human prostate cancer, cytostatic; immunostimulant; tumour; immunogen.
 Homo sapiens.
 NO200157190-A2.
 PN
 PD 09-AUG-2001.
 XX
 XX 05-FEB-2001; 2001NO-US04098.
 XX
 XX 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0568075.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUN-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0645936.
 PR 18-AUG-2000; 2000US-0638215.
 PR 02-OCT-2000; 2000US-0652719.
 PR 10-OCT-2000; 2000US-0679466.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 PA (HYSB-) HYSB INC.
 PI
 PI Tang YT, Liu C, Dymanc RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y,
 PI Zhao QA, Wang D, Wang J, Zhang Y, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Tang T, Wejman T, Goodrich R;
 DR N-PSDB; AAK51978.
 XX
 XX Nucleic acid encoding polypeptides with cytokine-like activities,
 XX useful in diagnosis and gene therapy -
 XX Claim 20; Page 3800-3801; 6221p; English.

CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAW78323-AAW80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

peptide therapy. The polypeptides have various cytokine-like activities, e.g., stem cell growth factor activity, haematopoietic regulatory activity, tissue growth factor activity, immunomodulatory activity and cell activity, and may be used to treat various diseases and conditions, e.g., cancer, leukemia, nervous system disorders, arthritis and inflammation. (Recorded for SEQ ID NO 2110 (AAK53581), 2111 (AAK53582) and 3666 (AAK80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication.)

Sequence 339 AA:

Query Match 100.0%; Score 116; DB 22; Length 339;
Best Local Similarity 100.0%; Pred. No. 1,4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHOQFFKIPILV 22
92 REVHPLATSHOQFFKIPILV 113

RESULT 9

AA001282 ID AA001282 standard; Protein: 339 AA.

AA001282;

04-OCT-2001 (first entry)

P789P amino acid sequence.

Human; prostate cancer; prostate-specific; diagnosis; vaccine;

cytotoxic; gene therapy; metastasis.

Homologous.

MO200151633-A2.

19-JUL-2001.

16-JAN-2001; 2001MO-US01574.

14-JAN-2001; 2000US-0483672.

(CORI-) CORIXA CORP.

Xu J, Dillon DC, Mutcham JL, Harlocker SL, Jiang Y, Reed SG;

Kalish MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skelty YAM;

Wang A, Mesinger MJ;

WPI; 2001-425873/45.

New polynucleotide encoding a prostate-specific protein, for

diagnosing, monitoring and treating prostate cancer in a patient and

for use in vaccines.

Claim 2; Page 510-512; 543pp; English.

The present invention describes polynucleotide sequences (I) which encode prostate-specific proteins (II) (I) and (II) have cytotoxic activity, and can be used in vaccine production and gene therapy. (I), (II), antibodies to (II), fusion proteins comprising (II), and isolated T cells prepared using (I) or (II) are used to treat cancer in a patient. (I) and the antibodies are also used in the detection of cancer in a patient. The cancer that is diagnosed or treated is particularly prostate cancer. The invention also describes polynucleotides or proteins encoded for monitoring the progression of cancer in a patient. (I) and (II) can also be used to improve diagnostic and therapeutic methods for prostate cancer. They can indicate the level of metastasis as well as the prostate volume. AA93357 to AA93394 and AA001115 to AA001119 represent polynucleotide and amino acid sequences used in the exemplification of the present invention.

Sequence 339 AA:

Query Match 100.0%; Score 116; DB 22; Length 339;
Best Local Similarity 100.0%; Pred. No. 1,4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHOQFFKIPILV 22
92 REVHPLATSHOQFFKIPILV 113

RESULT 10

ABG61813 ID ABG61813 standard; Protein: 339 AA.

ABG61813;

15-AUG-2002 (first entry)

Prostate cancer-associated protein #14.

Prostate cancer; prostate tumour tissue; human; mammal; cytotoxic.

Mammalia.

MO200230268-A2.

18-APR-2002.

12-OCT-2001; 2001MO-US32045.

13-OCT-2001; 2000US-0687576.

08-DEC-2001; 2000US-0733288.

24-JAN-2001; 2001US-2639572.

16-MAR-2001; 2001US-276791P.

16-MAR-2001; 2001US-276888P.

06-APR-2001; 2001US-281922P.

24-APR-2001; 2001US-286214P.

30-APR-2001; 2001US-0847046.

04-MAY-2001; 2001US-288598P.

(BOSB-) BOS BIOTECHNOLOGY INC.

Gish KC, Mack DH, Wilson KB, Afari D, Hovezi P;

WPI; 2002-471335/50.

N-PSDB; ABK92118.

Directing a prostate cancer-associated transcript in a cell in a

patient useful for diagnosing prostate cancer (PC) or screening

modulators of PC, by determining if prostate cancer-associated genes

are expressed in a prostate tissue.

Claim 27; Page 312; 436pp; English.

The present invention relates to methods of detecting a prostate cancer-associated transcript in a cell in a patient useful for diagnosing prostate cancer (PC) or screening modulators of PC, by determining if prostate cancer-associated genes are expressed in a prostate tissue. The method comprises contacting a biological sample from the patient with a selectively hybridized to a sequence that is at least 80% identical to them. The prostate cancer-associated polynucleotide sequences are differentially expressed in prostate tumour tissue or in prostate cancer and are derived from the tissues of various organisms such as humans or other mammals (e.g., mice, sheep and dogs). The method also includes identifying the prostate cancer-associated genes useful for diagnosing or treating prostate cancer, as well as for identifying modulators of prostate cancer or agents that inhibit prostate cancer. The nucleic acid sequences are particularly useful in gene therapy, as a vaccine or in antisense applications. ABG61800-ABG61944 represent prostate cancer-associated proteins.

SQ Sequence 339 AA; Score 116; DB 23; Length 339;
 Query Match 100.0%; Predicted No. 1,4e-10;
 Query Local Similarity 100.0%; Predicted No. 1,4e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 REVHPLATSHQGFYKIPILV 22
 DB 92 REVHPLATSHQGFYKIPILV 113
 RESULT 11
 ABB95387
 ID ABB95387 standard; Protein; 339 AA.
 AC ABB95387;
 XX
 DT 19-JUL-2002 (first entry)
 XX
 XX Human p789 protein SEQ ID NO 879.
 XX
 XX Human; cancer; prostate cancer; vaccine; cytosolic; immunostimulant;
 KW gene therapy.
 KM
 XX Homo sapiens.
 OS
 XX
 XX US2002022248-A1.
 XX
 XX 21-FEB-2002.
 XX
 PF 12-JAN-2001; 2001US-0759143.
 XX
 XX 25-FEB-1997; 97US-0806099.
 PR 01-AUG-1997; 97US-0904804.
 PR 06-FEB-1998; 98US-020956.
 PR 25-FEB-1998; 98US-023497.
 PR 14-JUL-1998; 98US-015453.
 PR 23-SEP-1998; 98US-0159812.
 PR 15-JAN-1999; 99US-0232149.
 PR 09-APR-1999; 99US-0289946.
 PR 13-JUL-1999; 99US-0352616.
 PR 12-NOV-1999; 99US-0439313.
 PR 15-JAN-2000; 2000US-0484876.
 PR 18-JAN-2000; 2000US-0536857.
 PR 27-MAR-2000; 2000US-0568100.
 PR 09-MAY-2000; 2000US-0570737.
 PR 12-MAY-2000; 2000US-0593793.
 PR 13-JUN-2000; 2000US-0605783.
 PR 27-JUN-2000; 2000US-0635215.
 PR 10-AUG-2000; 2000US-0638235.
 PR 25-SEP-2000; 2000US-0657236.
 PR 28-SEP-2000; 2000US-0657236.
 PR 02-OCT-2000; 2000US-0678426.
 PR 10-OCT-2000; 2000US-0686166.
 XX
 PA (XUJ/) XU J.
 PA (DILL) DILLON D. C.
 PA (MITC) MITCHELL J. L.
 PA (HACK) HACKER S. L.
 PA (KALOS) KALOS M. D.
 PA (FANG) FANGER G. R.
 PA (RETT) RETTER M. M.
 PA (STOL) STOLK J. A.
 PA (DAYC) DAY C. H.
 PA (VEDV) VEDVICK T. S.
 PA (HEND) HENDERSON R. A.
 PA (WANG) WANG A.
 PA (SKEI) SKEIKY Y. A. M.
 PA (HEPL) HEPLER W. T.
 PA (HEND) HENDERSON R. A.
 XX

PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD,
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
 PI Li SX, Wang A, Skeiky YAM, Hepler WT, Henderson RA;
 XX WPI, 2002-255649/30.
 DR
 PT New prostate-specific polynucleotides for diagnosing and treating
 PT disease, in particular prostate cancer, and as markers for the
 PT progression of cancer.
 XX
 XX Claim 2; SEQ ID NO 879; 87pp; English.
 CC The present invention provides prostate-specific coding sequences and
 CC their encoded proteins. These can be used in the diagnosis and treatment
 CC of cancer, particularly prostate cancer. The present sequence is a
 CC protein described in the invention.
 XX
 SQ Sequence 339 AA;
 Query Match 100.0%; Score 116; DB 23; Length 339;
 Query Local Similarity 100.0%; Predicted No. 1,4e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 REVHPLATSHQGFYKIPILV 22
 DB 92 REVHPLATSHQGFYKIPILV 113
 RESULT 12
 AAM79829
 ID AAM79829 standard; Protein; 374 AA.
 AC AAM79829;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 3475.
 XX
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; Leukemia;
 KW nervous system disorder; arthritis; inflammation.
 KM
 XX Homo sapiens.
 OS
 XX
 XX W0200157190-A2.
 XX
 PN 09-AUG-2001.
 XX
 PD 05-FEB-2001; 2001WO-US04098.
 XX
 PF 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-064936.
 PR 15-SEP-2000; 2000US-0653561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 PA (HYSE-) HYSEGO INC.
 XX
 XX Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y,
 PI Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW,
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI, 2001-476283/51.
 DR
 XX N-Psdb; AAM52962.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX
 PS Claim 20; Page 350; 6221pp; English.

CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAK481823-AAK60302) that exhibit activity eliciting to
 CC the production of polyclonal antibodies in the cell populations. The
 CC products of the polynucleotides are useful in gene therapy, vaccines or
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, hematopoietic regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and
 CC other diseases.
 CC Note: Records for SBO ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAK60020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.

CC Sequence 374 AA:

Query Match 100.0%; Score 116; DB 22; Length 374;
 Best Local Similarity 100.0%; Pred. No. 1, 5e-10;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGFYKPIPLV 22
 127 REVHPLATSHQGFYKPIPLV 148

RESULT 13

AAE02780 ID AAE02780 standard; Protein: 375 AA.

XX AAE02780;

XX 06-AUG-2001 (first entry)

XX Human six transmembrane epithelial antigen of prostate (STEAP)-1 protein.

XX Human; cytostatic; antiproliferative; vaccine; gene therapy;

XX six transmembrane epithelial antigen of the prostate-1; STEAP-1;

XX chromosome 7p22.3; cancer; prostate; colon; bladder; lung; ovarian;

XX pancreatic.

XX Homo sapiens.

XX Location/Qualifiers

FT Region 14..28 /label= Immungenic_peptide #1

FT Domain 70..91 /label= Transmembrane_domain #1

FT 91..99 /label= Transmembrane_domain #1

FT Region /label= HLA-A2_binding_peptide #2

FT Domain 114..136 /label= Transmembrane_domain #2

FT Region 158..166 /label= HLA-A2_binding_peptide #5

FT Domain 163..184 /label= Transmembrane_domain #3

FT Region 184..192 /label= HLA-A2_binding_peptide #1

FT Domain 219..241 /label= Transmembrane_domain #4

FT Domain 254..276 /label= Transmembrane_domain #5

FT Region 263..270 /label= HLA-A2_binding_peptide #3

FT Domain 292..313 /label= Transmembrane_domain #6

FT Region 302..310 /label= HLA-A2_binding_peptide #4

FT Msec-difference 339..340 /note= "Encoded by TTYGAGAT"

PN MO200140276-A2.

XX 07-JUN-2001.

XX 06-DEC-2000; 2000MO-US33040.

XX 06-DEC-1999; 99US-0455486.

XX (UROG-) UNOGENESIS INC.

XX Afar, DEH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC, Faris M,

XX Juckovaite A;

XX NPI: 2001-367804/38.

XX N-PSDB; AAD07067.

XX New STEAP (six transmembrane epithelial antigen of the prostate)

XX protein, expressed in human cancers, useful for detecting and treating

XX cancer -

XX Example 2; Fig 1A-1B, 187pp; English.

XX The present sequence is human six transmembrane epithelial antigen of

XX the prostate (STEAP)-1 protein of clone 10. STEAP-1 gene is located on

XX chromosome 7p22.3 and is used in gene therapy. Inhibiting the development

XX or progression of a cancer, especially colon, bladder, kidney, ovary

XX and prostate cancer, by administering a vaccine comprising a STEAP-1

XX expressing STEAP in a patient, comprises administering a vaccine

XX composition to the patient. Treating a patient with a cancer that

XX comprises administering to the patient a vector encoding single chain

XX and light chains of the monoclonal antibody that specifically binds to

XX sequence to the cancer cells and the encoded single chain monoclonal

XX Note: The present sequence is also shown in sequence listing of the

XX specification, but it lacks amino acid residues at its N-terminal end.

XX Sequence 375 AA:

Query Match 100.0%; Score 116; DB 22; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1, 5e-10;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REVHPLATSHQGFYKPIPLV 22
 92 REVHPLATSHQGFYKPIPLV 113

RESULT 14

AAV11840 ID AAV11840 standard; Protein: 95 AA.

XX AAV11840;

XX 18-JUN-1999 (first entry)

XX Human 5' EST secreted protein SBO ID No: 440.

XX Human; secreted protein; EST; expressed sequence tag; diagnosis;

XX forensic; gene therapy; chromosome mapping; signal peptide; prostate;

XX upstream regulatory sequence; cytokine activity; cell proliferation;

XX differentiation; hematopoietic regulation; tissue growth regulation;

XX reproductive hormone regulation; chemotactic; haemokinetic; haemostatic;

XX chromolytic; anti-inflammation; tumour inhibition.

OS Homo sapiens.

PN MO9306550-A2.

XX 11-FEB-1999.

PF	31-JUL-1998;	98NM0-1B01232.
XX		
PR	01-AUG-1997;	97US-0905144.
XX		
XX	(GEST) GENSET.	
PI	Duclet A, Dumas Milne Edwards J, Lacroix B;	
PI	WPI; 1999-133780/13.	
DR	N-Psdb; AAX40562.	
XX		
PT	New isolated prostate-derived nucleic acids - used to develop	
PT	product which may have cytokine, immune regulatory, haematopoiesis	
PT	regulating, anti-inflammatory or tumour inhibition activity	
PS	claim 34; Page 577; 675bp; English.	
XX		
CC	AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for	
CC	human secreted proteins expressed in prostate, and encode the proteins	
CC	given in AAY1716 to AAY1193 respectively. The proteins given represent	
CC	the signal peptide and an N-terminal fragment of a secreted protein. The	
CC	products may also be used for obtaining corresponding promoter	
CC	products. The proteins obtained may have cytokine activity, cell	
CC	proliferation and differentiation activity, haematopoiesis regulating	
CC	activity, tissue growth regulating activity, reproductive hormone	
CC	regulating activity, chemotactic/chemokinetic activity, haemostatic and	
CC	hemodynamic activity, receptor/ligand activity, anti-inflammatory	
CC	activity, tumour inhibition activity or other activities. The products	
CC	may also be used for obtaining corresponding promoter	
CC	sequences. The nucleic acids encoding the signal peptides can be used for	
CC	directing extracellular secretion of a polypeptide or the insertion of a	
CC	polypeptide into a membrane, or importing a polypeptide into a cell.	
XX		
XX	Sequence 95 AA;	
XX		
Query Match	83.6%; Score 97; DB 20; Length 95;	
Best Local Similarity	100.0%; Pred. No. 4; 2e-08;	
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 REVHPLATSHOQFFPKI 18	
DB	77 REVHPLATSHOQFFPKI 94	
RESULT 15		
ID	AAY12304	
XX	AAY12304 standard; Protein; 109 AA.	
AC	AAY12304;	
XX	17-JUN-1999 (Eitac entry)	
DE	Human 5' EST secreted protein SEQ ID NO.335.	
XX		
XX	Human; secreted protein; EST; expressed sequence tag; diagnosis;	
XX	forensic; gene therapy; chromosome mapping; signal peptide;	
XX	upstream regulatory sequence; cytokine activity; cell proliferation;	
XX	differentiation; haematopoiesis regulation; tissue growth regulation;	
XX	reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;	
XX	thrombolytic; anti-inflammatory; tumour inhibition.	
OS		
XX	Homo sapiens.	
PN	NO9906548-A2.	
XX		
XX	11-FEB-1999.	
XX		
PF	31-JUL-1998; 98NM0-1B01222.	
XX		
XX	01-AUG-1997; 97US-0905135.	

```

PA (GERT ) GENSER.
PI
XX Ductect A, Dumas Milne Edwards J, Lacroix B;
XX
XX MFI, 1999-15378/13.
XX
XX N-78DB, MAX1137.
XX
XX New nucleic acids encoding human secreted proteins - obtained from
XX cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
XX kidney, lung, umbilical cord, placenta and colon tissue
XX
XX Claim 27, Page 677, 824pp; English.
XX
XX AA10194 to AA11317 represent 5' expressed sequence tags (ESTs) for
XX human secreted proteins and encode the proteins given in AA11261 to
XX AA11264, respectively. The proteins given represent the signal peptide
XX and an N-terminal fragment of a secreted protein. The nucleic acid
XX sequences can be used for producing secreted human gene products. They
XX can also be used to develop products for diagnosis and therapy. The
XX proteins obtained may have cytokine activity, cell
XX proliferation/differentiation activity, hematopoietic regulating
XX activity, angiogenesis activity, immunomodulatory activity,
XX regulating activity, chemotactic/chemokinetic activity, haemostatic and
XX thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, tumour inhibition activity or other activities. The products
XX can be used in forensic, gene therapy and chromosome mapping procedures.
XX The sequences can also be used for obtaining corresponding promoter
XX sequences. The nucleic acids encoding the signal peptide can be used for
XX directing extracellular secretion of a polypeptide or the insertion of a
XX polypeptide into a membrane, or importing a polypeptide into a cell.
XX
XX Sequence 109 AA;
XX
XX Query Match 79.3%; Score 92; DB 20; Length 109;
XX Best Local Similarity 94.4%; Pred. No. 3.1e-07;
XX Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0.
XX
XX 1 REVHPLATSHQVFPYK 18
XX |||||
XX Db 92 REVHPLATSHQVFPYK 109

```


ID	CYF_PORPU	STANDARD;	PRT; 320 AA.
AC	PS1265;		
DT	01-OCT-1996 (rel. 34,	Created)	
FT	15-OCT-2001 (rel. 40,	Last sequence update)	
DR	15-OCT-2001 (rel. 40,	Last annotation update)	
OK	Apocycochrome F precursor.		
GN	PETA.		
OG	Porphyra purpurea.		
OC	Chloroplast.		
CC	Bakryocae; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.		
CC	NBOL_Taxid=2787;		
CC	SEQUENCE FROM N.A.		
CC	SRAIN=avonport/.		
CC	Retch M.E., Muniholland J.;		
CC	Complete nucleotide sequence of the Porphyra purpurea chloroplast genome."		
CC	Plant Mol. Biol. Rep. 13:333-335(1995).		
CC	- FUNCTION: TRANSLATES PROTONS ACROSS THE THYLAKOID MEMBRANE AND		
CC	REMOVES EXCESS IRON FROM THE CYTOCHROME C FAMILY. THIS PROTEIN PASSES		
CC	THEM TO PLASTOCYANIN. THIS FUNCTION IS VERY SIMILAR TO THAT OF		
CC	MITOCHONDRIAL CYTOCHROME C1.		
CC	- SUBUNIT: MEMBER OF THE CYTOCHROME b6/f COMPLEX INCLUDING		
CC	CYTOCHROME b6, CYTOCHROME f AND PROBABLY AN IRON SULFUR PROTEIN.		
CC	- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane (probable).		
CC	- SIMILARITY: BELONGS TO THE CYTOCHROME c FAMILY.		
CC	This Swiss-Prot entry is copyrighted. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL consortium -		
CC	the European Bioinformatics Institute. There are no restrictions on ways		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/)		
CC	or send an email to license@sb-sib.ch .		
CC	PMID: 1198004, PMO0351.1).		
DR	HSPD: P36438, IHCZ.		
DR	InterPro: IPRO02325; APOCYE.F.		
DR	InterPro: IPRO0345; CYCt Heme bind.		
DR	Pfam: PF01333; Apocycochrome_F_1.		
DR	PRINTS: PR00610; CYTOCHROME-F.		
DR	PROSITE: PS00190; CYTOCHROME_C_1.		
KW	Eosinophilic; Hemoglobin; Hemoglobins; Hemoproteins; Thylakoids;		
KW	Transmembrane; Phycocyanin II; Phycocyanin II; Transist peptide;		
KW	Transmembrane		
FT	TRANSIT 1 36 CHLOROPLAST (BY SIMILARITY).		
FT	CHAIN 37 320 APOCYTOCHROME F.		
FT	BINDING 57 57 HEME (COVALENT) (PROBABLE).		
FT	BINDING 60 60 HEME (COVALENT) (PROBABLE).		
FT	METVAL 61 61 IRON (HEME AXIAL LIGAND) (PROBABLE).		
FT	TRANSMEM 268 305 POTENTIAL.		
SO	SEDOHENSE 320 AA, 35012 MW, CSN1059753169CT CRC64,		
QY	Query Match 40.3%; Score 47.5; DB 1; Length 320;		
Db	Beet Local Similarity 40.7%; Pred. No. 2.8;		
	Matches 11; Conservative 4; Mismatches 7; Indels 5; Gaps 1.		
	160 REIPIFLSDPDKMDKMFIRIPV 166		
RESULT 5			
SYFA ARATH	STANDARD;	PRT; 474 AA.	
AC	GPT034; 2001 (rel. 40,	Created)	
DT	15-OCT-2001 (rel. 40,	Last sequence update)	
DR	15-OCT-2001 (rel. 40,	Last annotation update)	
DE	Probable phenylalanine-tRNA synthetase alpha chain (EC 6.1.1.20)		
DE	(Phenylalanyl-tRNA ligase alpha chain) (PheRS).		
DE	AT4G33280 OR T328.180.		

OS Arabidopsis thaliana (Mouse-ear cress).
CC Bakersyca; Viridiplantae; Streptophyta; Embryophyta; Tracheophytes;
CC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicot; Rosidae;
CC Malvaceae; Malvales; Brassicales; Brassicaceae; Brassicaceae;
CC NCBI TaxID=3702.
CC [1]_TaxID=3702.
RP SEQUENCE FROM N.A.
RP STRAIN=ecv, Columbia;
RX MEDLINE=20083489; PubMed=16517198;
RA Meyer K.F.X., Schueller C., Wandt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoefer A., Stiekema W., Enkin K., Duerren M.,
RA Kottmann J., Bressanini M., de Simon V., Obermair B., Meiche R.,
RA Weichelt J., Gornall M., de Simon V., Obermair B., Meiche R.,
RA Kottmann J., Bressanini M., de Simon V., Obermair B., Meiche R.,
RA Reichert B., Portelle D., Perez-Alonso M., Boutin M., Bancroft I.,
RA Vos P., Hobeisel U., Zimmermann W., Medler H., Riddle P.,
RA Langham S.-A., McCallum B., Bihlma I., Robben J.,
RA Van der Schueren J., Grynopowicz B., Chuang Y.-J., Vandewussche F.,
RA Braeken M., Welfelmis I., Voet M., Baetlemis I., Aert R., Devoe E.,
RA Holter E., Brandt A., Peters S., van Steveninck M., Dikite W.,
RA Moosman P., Klein Lamborch R., Rose M., Hunt J., Koester P.,
RA Bernreiter S., Hempelhorst S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buyssebaert C., Gielens J., Vialletoen R., De Clercq R.,
RA Van Montagu M., Rogers U., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Key M., Leonard N., McIay K.,
RA Pectet A., Ralmandem M.A., Lyne M., James V., Reclman S.,
RA Doe S., De Haese M., Marais A.C., M. G. M., Muller-Aner S.,
RA Gabel C., Fuchs M., Fattmann B., Granderath K., Danner D., Herzl A.,
RA Neumann S., Aguilou A., Chaleid D., Ligouri R., Plevander R.,
RA Messene O., Ousley F., Clapaud G., Wendlein A., Faltner F.,
RA Schnabl S., Hiller R., Schmidt W., Lechery A., Aubourg S.,
RA Chetoui F., Cooke R., Berger G., Montfort A., Canaberta E.,
RA Gidbom T., Weber M., Pandolfi M., Bagheri S., Teyssie D., Torres A.,
RA Heijlen L., Schwartz S., Scholler P., Heber S., France P., Bialek C.,
RA Frilman D., Haese D., Lemke C., Mews H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Haerem K.,
RA Parnell L., Debia N., Gnoj L., Schutz K., Huang E., Spigel J.,
RA Selson M., Murray J., Shee P., Cordes M., Abu-Threiden U.,
RA Smetzing T., Kalicki J., Graves T., Hatton G., Edwards J.,
RA Lattelle P., Courtney L., Cloud M., Abbott A., Scott R., Johnson D.,
RA Kimm J., Benito L., Madis E., Dantes M., Pajot K., Hillier L.,
RA Nelson J., Splich U., Ryan E., Andrews S., Gatsel C., Layman D.,
RA Du H., Ali U., Berghoff A., Jones K., Dione K., Cotton M., Joshi C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
RA Ma P., Zhong Y., Preston R., Vill D., Shaker M., Mateo A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman U., Till S.,
RA Grant S., Shony N., Nagesawa A., Hamed A., Lodi M., Johnson A.,
RA Gen E., Maria M., Matsumura A., Meece M., Meece M.,
RA Nelson J., Splich U., Ryan E., Andrews S., Gatsel C., Layman D.,
RT Nature 402:769-771(1999).
CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA (Phe) = AMP +
CC diphosphate + L-phenylalanyl-tRNA(Phe).
CC -1- SUBSTRATE: RETANES OF TWO ALPHA AND TWO BETA CHAINS (BY
CC SIMILARITY). LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC PHE-TRNA SYNTHETASE ALPHA CHAIN SUPPLEMENT 2.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is not republished
CC without the EMBL/EBI logo and the EMBL/EBI accession number (EMBL/EBI
CC entries require a license agreement (see [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
CC or send an email to license@ebi.ac.uk).
CC -----
DB EMBL, AL050351, CA94363.1, AL1.SFO.
DB EMBL, AL161594, CAB80591.1, AL1.SFO.

[illegible]

KM	Hypothetical protein; Complete proteome.
XW	SEQUENCE 737 AA; 82605 MW; 567BE9A18B5391F1F CACG64;
Qy	5 HPLATSIHQYEFKXP 19 Db 389 HPMAEHRHQMTELP 403
DR	RESULT 7
ID	CYTA_SACKL STANDARD; PRT; 1839 AA.
AC	P23466;
DT	01-NOV-1991 (Rel. 20). Created)
DY	01-NOV-1991 (Rel. 20). Last sequence update)
DE	15-JUN-2002 (Rel. 41). Last annotation update)
DF	AdeNyvate cyclase [EC 4.6.1.1] (ATP pyrophosphate-lyase) [Adenylyl transferase].
GN	Saccharomyces kluyveri (Yeast).
GC	Osccharomyces kluyveri (Yeast).
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX	NCEI_TaxID=4934; 11
RN	SEQUENCE FROM N.A. MEDLINE=61321718, PubMedId=1665503; RA Young D., O'Neill K., Break D., Wiegler M.; RL "The adenylyl cyclase-encoding gene from Saccharomyces kluyveri." J Gene 102:129-132(1991).
RT	- FUNCTION: PLAYS ESSENTIAL ROLE IN REGULATION OF CELLULAR METABOLISM BY CATALYZING THE SYNTHESIS OF A SECOND MESSENGER,
CC	- CATALYTIC ACTIVITY: ATP = 3'-5'-cyclic AMP + diphosphate.
CC	- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-3 FAMILY.
CC	- SIMILARITY: CONTAINS 2 LEUCINE-RICH REPEATS (LRR).
CC	- SIMILARITY: CONTAINS 1 P2C2-LIKE DOMAIN.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. The content is provided as-is without warranty of any kind. In no way shall the EBI or its staff be liable for damages arising from the use of this modified and this statement is not removed. Usage by end user for commercial purposes requires a license agreement (see http://www.ebi.ac.uk/announcements send an email to licenses@ebi.ac.uk).
CC	-- or send an email to licenses@ebi.ac.uk --
DR	EMBL; X56042; CAJ39513.1; -
DR	PIR; S1464; ORBYX
DR	FIR; QLI145; QLI145
DR	IPIPO01051; IPIPO01051; G_cyclase.
DR	InterPro; IPR001161; LRR
DR	InterPro; IPR003592; LRR_out
DR	InterPro; IPR003591; LRR_TYP
DR	InterPro; IPR001932; P2C2_like
DR	InterPro; IPR000159; RA_domain
DR	Pfam; PF00211; gnan_yacc_CYC_1
DR	Pfam; PF00460; LRR_1
DR	PRINTS; PR00019; LEINCHPT.
DR	SMART; SM00044; CYCC; 1.
DR	SMART; SM00312; P2C2; 1.
DR	SMART; SM00314; RA_1
DR	POSITIVE; PS51255; GUANYLATE_CYCLASE_2_1
KT	Repeat 653 Leucine-rich LRR 1
FT	REPEAT 659 679 LRR 2
FT	REPEAT 680 702 LRR 3
FT	REPEAT 703 724 LRR 4
FT	REPEAT 726 748 LRR 5
Tt	

FT REPEAT 749 771 LRR 6.
 FT REPEAT 772 793 LRR 7.
 FT REPEAT 824 836 LRR 8.
 FT REPEAT 834 857 LRR 9.
 FT REPEAT 859 879 LRR 10.
 FT REPEAT 880 903 LRR 11.
 FT REPEAT 905 936 LRR 12.
 FT REPEAT 927 949 LRR 13.
 FT REPEAT 951 974 LRR 14.
 FT REPEAT 1026 1058 LRR 15.
 FT REPEAT 1059 1072 LRR 16.
 FT REPEAT 1074 1096 LRR 17.
 FT REPEAT 1101 1124 LRR 18.
 FT REPEAT 1135 1160 LRR 19.
 FT REPEAT 1185 1440 LRR 20.
 FT DOMAIN 1441 1839 P2/C-LIKE.
 FT SEQUENCE 1839 AA; 206695 MW; 86469BCMF7J3JCB CRC64;

Query Match 39.7%; Score 46; DB 1; Length 1839;
 Best Local Similarity 66.7%; Pred. No. 35;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 VHPPLATSHOXY 14
 DB 589 VHPPLATSHOXY 600

RESULT 8
 FRAP MOUSE STANDARD; PRT; 2549 AA.
 AC GJUN19; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE FRAP-rapamycin associated protein (FRAP).
 OS FRAP1 OR FRAP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;

RA STRAIN=BAALB/C;
 RC Blk/roky V.; Mock B.;
 RL "Positional cloning of mouse plasmacytoma susceptibility gene."
 RT Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: ACTS AS THE TARGET FOR THE CELL-CYCLE ARREST AND
 CC IMMUNOSUPPRESSIVE EFFECTS OF THE FRAP12-RAPAMYCIN COMPLEX.
 CC -1- SIMILARITY: HOMOLOG TO HEP REPEATS.
 CC -1- SIMILARITY: BELONGS TO THE KINASES FAMILY.

CC This SMS-S-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.1sb-sib.ch/announce/>
 CC or <http://www.1sb-sib.ch/>)

CC EMBL; AF153838; AAF73196.1; -;
 DR HSP; P23345; 1SPF.
 DR MGD; MGI:1928394; Fpzd1.
 DR InterPro; IPR003151; FRT.
 DR InterPro; IPR003152; FRT.
 DR InterPro; IPR004043; P13_P14_kinase.
 DR Pfam; PF00454; P13_P14_kinase; 1.
 DR Pfam; PF02259; FRT; 1.
 DR Pfam; PF02260; FRT; 1.
 DR SMART; SM00146; P13K; 1.
 DR PROSITE; PS00915; P13_KINASE_1; 1.

DR PROSITE; PS00916; P13_KINASE_2; 1.
 DR PROSITE; PS00920; P13_KINASE_3; 1.
 DR PROSITE; PS00977; HEAT_REPEAT; FALSE_NEG.
 DR TrEMBL; P13K; 1.
 FT REPEAT 163 185 HEAT 1.
 FT REPEAT 650 688 HEAT 2.
 FT REPEAT 689 697 HEAT 3.
 FT REPEAT 859 897 HEAT 4.
 FT REPEAT 988 1025 HEAT 5.
 FT REPEAT 1069 1106 HEAT 6.
 FT REPEAT 1109 1148 HEAT 7.
 FT REPEAT 1159 1185 HEAT 8.
 FT REPEAT 1931 1970 HEAT 9.
 FT DOMAIN 2182 2549 P13K/P14K.
 FT SEQUENCE 2549 AA; 288734 MW; C02636E6F61BC3 CRC64;

Query Match 38.8%; Score 45; DB 1; Length 2549;
 Best Local Similarity 50.8%; Pred. No. 73;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 REVHPLATSHOXY 16
 DB 1683 RQDHPPLATSHOXY 1698

RESULT 9
 SPRT SEMI_PIG STANDARD; PRT; 137 AA.
 AC Q26920;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Seminal plasma sperm motility inhibitor precursor.
 OS SEMI_PIG.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 NCBI_TaxID=96823;

RA TISSUE=seminal vesicle;
 RC MEDLINE; 953194; PubMed/953190;
 RA Oada T.; Gagnon C.; H., Furlan Y., Wada K., Satoh M., Satoh M.,
 RT "Cloning of boar SEMI gene which is expressed specifically in seminal
 RT vesicle and codes for a sperm motility inhibitor protein".
 RL FEMS Lett. 368:420-424(1995).
 CC -1- FUNCTION: INHIBITOR OF SPERM MOTILITY.
 CC -1- SUBCELLULAR LOCATION: SPERM.
 CC -1- TISSUE SPECIFICITY: SEMINAL PLASMA OR SPERM.
 CC -1- SIMILARITY: BELONGS TO THE SPERMADHESIN FAMILY.

CC This SMS-S-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.1sb-sib.ch/announce/>
 CC or <http://www.1sb-sib.ch/>)

CC EMBL; S80568; AAB35000.2; -;
 DR HSP; P23392; 1SPF.
 DR InterPro; IPR000859; CUB_domain.
 DR InterPro; IPR001024; Spermadhesin.
 DR Pfam; PF00311; CUB; 1.
 DR Pfam; SM00147; CUB; 1.
 DR PROSITE; PS00985; SPERMADHESIN_1; 1.
 DR PROSITE; PS00986; SPERMADHESIN_2; FALSE_NEG.
 DR PROSITE; PS01180; CUB; 1.
 DR Heparin-binding; Glycoprotein; Fertilization; Signal.
 FT SIGNAL 22 21
 FT CHAIN 137
 FT SEMINAL PLASMA SPERM MOTILITY INHIBITOR.


```

FT DOMAIN          68      780
FT TRANSMEM       781      824
FT SIGNAL         825      932
FT CDS            933     1056
FT CARBOHYD       1057     1155
FT CARBOHYD       1156     1253
FT CARBOHYD       1254     1351
FT CARBOHYD       1352     1449
FT CARBOHYD       1450     1547
FT CARBOHYD       1548     1645
FT CARBOHYD       1646     1743
FT CARBOHYD       1744     1841
FT CARBOHYD       1842     1939
FT CARBOHYD       1940     2037
FT CARBOHYD       2038     2135
FT CARBOHYD       2136     2233
FT CARBOHYD       2234     2331
FT CARBOHYD       2332     2429
FT CARBOHYD       2430     2527
FT CARBOHYD       2528     2625
FT CARBOHYD       2626     2723
FT CARBOHYD       2724     2821
FT CARBOHYD       2822     2919
FT CARBOHYD       2920     3017
FT CARBOHYD       3018     3115
FT CARBOHYD       3116     3213
FT CARBOHYD       3214     3311
FT CARBOHYD       3312     3409
FT CARBOHYD       3410     3507
FT CARBOHYD       3508     3605
FT CARBOHYD       3606     3703
FT CARBOHYD       3704     3801
FT CARBOHYD       3802     3900
FT CARBOHYD       3901     3998
FT CARBOHYD       3999     4096
FT CARBOHYD       4097     4194
FT CARBOHYD       4195     4292
FT CARBOHYD       4293     4390
FT CARBOHYD       4391     4488
FT CARBOHYD       4489     4586
FT CARBOHYD       4587     4684
FT CARBOHYD       4685     4782
FT CARBOHYD       4783     4880
FT CARBOHYD       4881     4978
FT CARBOHYD       4979     5076
FT CARBOHYD       5077     5173
FT CARBOHYD       5174     5271
FT CARBOHYD       5272     5369
FT CARBOHYD       5370     5467
FT CARBOHYD       5468     5564
FT CARBOHYD       5565     5662
FT CARBOHYD       5663     5759
FT CARBOHYD       5760     5857
FT CARBOHYD       5858     5954
FT CARBOHYD       5955     6052
FT CARBOHYD       6053     6149
FT CARBOHYD       6150     6247
FT CARBOHYD       6248     6344
FT CARBOHYD       6345     6442
FT CARBOHYD       6443     6539
FT CARBOHYD       6540     6637
FT CARBOHYD       6638     6734
FT CARBOHYD       6735     6832
FT CARBOHYD       6833     6929
FT CARBOHYD       6930     7027
FT CARBOHYD       7028     7124
FT CARBOHYD       7125     7222
FT CARBOHYD       7223     7319
FT CARBOHYD       7320     7417
FT CARBOHYD       7418     7515
FT CARBOHYD       7516     7613
FT CARBOHYD       7614     7711
FT CARBOHYD       7712     7808
FT CARBOHYD       7809     7906
FT CARBOHYD       7907     8003
FT CARBOHYD       8004     8101
FT CARBOHYD       8102     8200
FT CARBOHYD       8201     8297
FT CARBOHYD       8298     8395
FT CARBOHYD       8396     8493
FT CARBOHYD       8494     8590
FT CARBOHYD       8591     8688
FT CARBOHYD       8689     8785
FT CARBOHYD       8786     8883
FT CARBOHYD       8884     8980
FT CARBOHYD       8981     9078
FT CARBOHYD       9079     9175
FT CARBOHYD       9176     9272
FT CARBOHYD       9273     9369
FT CARBOHYD       9370     9467
FT CARBOHYD       9468     9564
FT CARBOHYD       9565     9662
FT CARBOHYD       9663     9759
FT CARBOHYD       9760     9857
FT CARBOHYD       9858     9954
FT CARBOHYD       9955     10052
FT CARBOHYD       10053     10150
FT CARBOHYD       10151     10247
FT CARBOHYD       10248     10345
FT CARBOHYD       10346     10443
FT CARBOHYD       10444     10540
FT CARBOHYD       10541     10638
FT CARBOHYD       10639     10735
FT CARBOHYD       10736     10833
FT CARBOHYD       10834     10930
FT CARBOHYD       10931     11028
FT CARBOHYD       11029     11125
FT CARBOHYD       11126     11222
FT CARBOHYD       11223     11319
FT CARBOHYD       11320     11417
FT CARBOHYD       11418     11514
FT CARBOHYD       11515     11612
FT CARBOHYD       11613     11709
FT CARBOHYD       11710     11806
FT CARBOHYD       11807     11903
FT CARBOHYD       11904     12001
FT CARBOHYD       12002     12098
FT CARBOHYD       12099     12195
FT CARBOHYD       12196     12292
FT CARBOHYD       12293     12390
FT CARBOHYD       12391     12487
FT CARBOHYD       12488     12584
FT CARBOHYD       12585     12682
FT CARBOHYD       12683     12779
FT CARBOHYD       12780     12876
FT CARBOHYD       12877     12973
FT CARBOHYD       12974     13070
FT CARBOHYD       13071     13168
FT CARBOHYD       13169     13264
FT CARBOHYD       13265     13362
FT CARBOHYD       13363     13459
FT CARBOHYD       13460     13556
FT CARBOHYD       13557     13653
FT CARBOHYD       13654     13750
FT CARBOHYD       13751     13848
FT CARBOHYD       13849     13945
FT CARBOHYD       13946     14042
FT CARBOHYD       14043     14139
FT CARBOHYD       14140     14237
FT CARBOHYD       14238     14334
FT CARBOHYD       14335     14431
FT CARBOHYD       14432     14528
FT CARBOHYD       14529     14625
FT CARBOHYD       14626     14722
FT CARBOHYD       14723     14819
FT CARBOHYD       14820     14916
FT CARBOHYD       14917     15014
FT CARBOHYD       15015     15110
FT CARBOHYD       15111     15207
FT CARBOHYD       15208     15304
FT CARBOHYD       15305     15401
FT CARBOHYD       15402     15498
FT CARBOHYD       15499     15595
FT CARBOHYD       15596     15692
FT CARBOHYD       15693     15789
FT CARBOHYD       15790     15886
FT CARBOHYD       15887     15983
FT CARBOHYD       15984     16080
FT CARBOHYD       16081     16177
FT CARBOHYD       16178     16274
FT CARBOHYD       16275     16370
FT CARBOHYD       16371     16467
FT CARBOHYD       16468     16564
FT CARBOHYD       16565     16661
FT CARBOHYD       16662     16758
FT CARBOHYD       16759     16855
FT CARBOHYD       16856     16951
FT CARBOHYD       16952     17048
FT CARBOHYD       17049     17145
FT CARBOHYD       17146     17242
FT CARBOHYD       17243     17339
FT CARBOHYD       17340     17436
FT CARBOHYD       17437     17533
FT CARBOHYD       17534     17630
FT CARBOHYD       17631     17726
FT CARBOHYD       17727     17823
FT CARBOHYD       17824     17920
FT CARBOHYD       17921     18017
FT CARBOHYD       18018     18114
FT CARBOHYD       18115     18210
FT CARBOHYD       18211     18307
FT CARBOHYD       18308     18404
FT CARBOHYD       18405     18501
FT CARBOHYD       18502     18598
FT CARBOHYD       18599     18695
FT CARBOHYD       18696     18791
FT CARBOHYD       18792     18888
FT CARBOHYD       18889     18985
FT CARBOHYD       18986     19081
FT CARBOHYD       19082     19178
FT CARBOHYD       19179     19275
FT CARBOHYD       19276     19371
FT CARBOHYD       19372     19468
FT CARBOHYD       19469     19565
FT CARBOHYD       19566     19661
FT CARBOHYD       19662     19758
FT CARBOHYD       19759     19855
FT CARBOHYD       19856     19951
FT CARBOHYD       19952     20048
FT CARBOHYD       20049     20145
FT CARBOHYD       20146     20242
FT CARBOHYD       20243     20339
FT CARBOHYD       20340     20436
FT CARBOHYD       20437     20533
FT CARBOHYD       20534     20630
FT CARBOHYD       20631     20726
FT CARBOHYD       20727     20823
FT CARBOHYD       20824     20920
FT CARBOHYD       20921     21017
FT CARBOHYD       21018     21114
FT CARBOHYD       21115     21210
FT CARBOHYD       21211     21307
FT CARBOHYD       21308     21404
FT CARBOHYD       21405     21501
FT CARBOHYD       21502     21598
FT CAR
```

```

QY      1 REVHPLATSHOQYP 15
Db      659 RELVEPCTNKHRYF 673

RESULT 15
YHVS_ECOLI ID_YHVS_ECOLI STANDARD; PRT; 326 AA.
AC P39370;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE HYPOTHEtical protein yHvs precursor.
PR 1988, ORNL, YHVS_ECOLI
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Echerichia.
ON NCBI_TaxID=562;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K13 / MG1655;
RX MEDLINE=9534362; PubMed=7510040;
RX BLAST=9534362; PubMed=7510040;
RX Blumberg P.R., Plimmet G. III, Sotia H.J., Daniels D.L.,
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RL region from 92.8 through 100 minutes.";
RL Nucleic Acids Res. 23:2105-2119(1995).

CC CC This SMS-S-PROT entry is copy-right. It is produced through a collaboration
CC between the Swiss Institute of Biomedicine and the EMBL collaborator
CC The European Bioinformatics Institute, whose content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ldb-eb.ch/announce/ or
CC or send an email to licen@esl.bdb-eb.ch).

-----
DR EMBL; U14003; AAA97205.1; "-"
DR EMBL; AE000501; AAC7265.1; "-"
DR Ecocode; EC12561.yHVS.DU9703.
DR InterPro; IPR005181.1 DU9703.
DR Pfam; PF01423.1 yHVS.DU9703.
DR Hypothetical; yHVS.DU9703.Signal; Complete proteome.
FT SIGNAL 1 21 POTENTIAL
FT CHAIN 22 326 HYPOTHEtical PROTEIN YHVS.
SQ SEQUENCE 326 AA; 36678 MW; 6E40B170E7BDAD CIGC64;

Query Match
Best Local Similarity 81.8%; Score 43.5; DB 1; Length 326;
Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1.

OY 5 HPATSHD-QY 14
Db 80 HPATSHD-QY 90

```


GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 16:49:57 / Search time 14.8372 Seconds

(without alignments) 87.073 Million cell updates/sec

Title: US-10-010-667a-19

Sequence: 1 REVHPLATSHOQFKPIILV 22

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 237916 seqs, 58723674 residues

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Result No.	Query	Match	Length	DB ID	Description
1	116	100.0	22	9	US-10-011-095-19
2	116	100.0	22	9	US-10-010-667a-19
3	116	100.0	104	10	US-09-864-761-47234
4	116	100.0	267	10	US-09-747-835a-50
5	116	100.0	267	9	US-10-171-862873
6	116	100.0	339	9	US-09-402-520-11
7	116	100.0	339	9	US-09-895-793-879
8	116	100.0	339	9	US-09-895-814-879
9	116	100.0	339	9	US-10-011-095-2
10	116	100.0	339	9	US-10-010-667a-2
11	116	100.0	339	10	US-09-799-143-879
12	116	100.0	339	10	US-09-432-837-879
13	116	100.0	339	10	US-09-402-520-11
14	116	100.0	339	10	US-09-747-835a-15
15	116	100.0	339	10	US-09-747-835a-15
16	116	100.0	339	10	US-09-747-835a-15
17	116	100.0	339	10	US-09-747-835a-15
18	116	100.0	339	10	US-09-747-835a-15
19	116	100.0	339	10	US-09-747-835a-15
20	116	100.0	339	10	US-09-747-835a-15
21	116	100.0	339	10	US-09-747-835a-15
22	116	100.0	339	10	US-09-747-835a-15
23	116	100.0	339	10	US-09-747-835a-15
24	116	100.0	339	10	US-09-747-835a-15
25	116	100.0	339	10	US-09-747-835a-15
26	116	100.0	339	10	US-09-747-835a-15
27	116	100.0	339	10	US-09-747-835a-15
28	116	100.0	339	10	US-09-747-835a-15
29	116	100.0	339	10	US-09-747-835a-15
30	116	100.0	339	10	US-09-747-835a-15
31	116	100.0	339	10	US-09-747-835a-15
32	116	100.0	339	10	US-09-747-835a-15
33	116	100.0	339	10	US-09-747-835a-15
34	116	100.0	339	10	US-09-747-835a-15
35	116	100.0	339	10	US-09-747-835a-15
36	116	100.0	339	10	US-09-747-835a-15
37	116	100.0	339	10	US-09-747-835a-15
38	116	100.0	339	10	US-09-747-835a-15
39	116	100.0	339	10	US-09-747-835a-15
40	116	100.0	339	10	US-09-747-835a-15
41	116	100.0	339	10	US-09-747-835a-15
42	116	100.0	339	10	US-09-747-835a-15
43	116	100.0	339	10	US-09-747-835a-15
44	116	100.0	339	10	US-09-747-835a-15
45	116	100.0	339	10	US-09-747-835a-15

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

20	42	36.2	2549	9	US-09-950-634-3	Sequence 3, Appl1
21	41	35.3	469	10	US-09-865-282-10	Sequence 10, Appl1
22	40	34.5	202	12	US-10-127-913-2	Sequence 12, Appl1
23	40	34.5	359	9	US-10-055-364-48	Sequence 48, Appl1
24	40	34.5	430	10	US-09-799-777-62	Sequence 62, Appl1
25	40	34.5	437	9	US-09-738-626-4256	Sequence 426, Appl1
26	40	34.5	437	9	US-09-746-650a-2	Sequence 2, Appl1
27	40	34.5	437	10	US-09-919-932-2	Sequence 344, Appl1
28	40	34.5	1336	9	US-10-043-487-34	Sequence 344, Appl1
29	40	34.5	162	10	US-09-864-761-4491	Sequence 4491, Appl1
30	39	33.6	339	12	US-10-067-291-1	Sequence 1, Appl1
31	39	33.6	353	9	US-10-007-343-2	Sequence 2, Appl1
32	39	33.6	636	9	US-09-738-626-5590	Sequence 5590, Appl1
33	39	33.6	777	9	US-09-945-901-16	Sequence 16, Appl1
34	39	33.6	777	9	US-10-007-747-16	Sequence 16, Appl1
35	39	33.6	827	9	US-10-038-931-15	Sequence 15, Appl1
36	39	33.6	854	9	US-10-007-747-12	Sequence 12, Appl1
37	39	33.6	854	9	US-10-038-931-32	Sequence 32, Appl1
38	39	33.6	870	9	US-09-945-901-30	Sequence 30, Appl1
39	39	33.6	870	9	US-10-007-747-30	Sequence 30, Appl1
40	39	33.6	870	9	US-10-038-931-30	Sequence 30, Appl1
41	39	33.6	875	9	US-09-945-901-28	Sequence 28, Appl1
42	39	33.6	875	9	US-10-007-747-28	Sequence 28, Appl1

ALIGNMENTS

```

RESULT 1
US-10-011-095-19
Sequence 19, Application US/10011095
US-10-011-095-19
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Leong, Rene S.
APPLICANT: Saittan, Arthur B.
APPLICANT: Saittan, Douglas C.
APPLICANT: Saittan, Steve Campbell
APPLICANT: Saittan, Steve Campbell
FILE REFERENCE: 511582001610
CURRENT APPLICATION NUMBER: US/10/011.095
PRIOR FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: 09/323,873
PRIOR FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: 60/087,520
PRIOR FILING DATE: 1998-06-01
PRIOR FILING DATE: 1998-06-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 19
LENGTH: 22
TYPE: PRT
ORIGIN: Artificial Sequence
OTHER INFORMATION: STAMP-1 peptide
US-10-011-095-19

Query Match      100.0% Score 116; DB 9; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.4e-12;
Matches 22; Conservative 0; Mismatches 0; Gaps 0;
Cy 1 REVHPLATSHOQFKPIILV 22
Db 1 REVHPLATSHOQFKPIILV 22

RESULT 2
US-10-010-667a-19

```

```

Sequence 19, Application US/10010667a
Publication No. US20030055217A1
GENERAL INFORMATION:
APPLICANT: Aree, Daniel S.
INVENTOR: Aree, Daniel S.
APPLICANT: Raitano, Arthur B.
APPLICANT: Leong, Kahm
APPLICANT: Saffran, Douglas C.
APPLICANT: Mitchell, Steve Chappell
TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
CURRENT FILING DATE: 2001-01-30
CURRENT FILING DATE: 2001-01-30
CURRENT FILING DATE: 2001-12-06
PRIOR FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: 09/223,873
PRIOR FILING DATE: 1998-06-01
PRIOR APPLICATION NUMBER: 60/087,520
PRIOR FILING DATE: 1998-06-01
PRIOR APPLICATION NUMBER: 60/091,183
NUMBER OF SEQ ID NOS: 32
SOFTWARE: ParsSeq for Windows Version 4.0
SEQ ID NO: 19
LENGTH: 22
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: STAMP-1 peptide
US-10-010-667a-19

```

```

Query Match          100.0%; Score 116; DB 9; Length 22;
Best Local Similarity 100.0%; Pred. No. 7,4e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 REVHPLATSHQYFYKIPILV 22
DB      1 REVHPLATSHQYFYKIPILV 22

```

```

RESULT 3
US-09-864-761-47234
Sequence 47234, Application US/09864761
Publication No. US2002018783A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
INVENTOR: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wenhang
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
CURRENT FILING DATE: 2001-05-23
CURRENT FILING DATE: 2001-09/864,761
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/271,456
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US 60/322,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30

```

```

PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine ver8. 1.1
SEQ ID NO: 1224
LENGTH: 104
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC005053.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.87
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.59
OTHER INFORMATION: SWISSPROT HIT: O67305, EVALU8 5,60e+00
OTHER INFORMATION: EST_HUMAN HIT: BB875216.1, EVALU8 2,00e-54
US-09-864-761-47234

```

```

Query Match          100.0%; Score 116; DB 10; Length 104;
Best Local Similarity 100.0%; Pred. No. 7,4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 REVHPLATSHQYFYKIPILV 22
DB      59 REVHPLATSHQYFYKIPILV 80

```

```

RESULT 4
US-09-747-835A-50
Sequence 50, Application US/09747835A
Patent No. US20020146692A1
GENERAL INFORMATION:
APPLICANT: Yamazaki, Victoria
INVENTOR: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Zhou, Ding
APPLICANT: Zhou, Daniel
APPLICANT: Zhang, Jie
APPLICANT: Ren, Peiyan
APPLICANT: Asundi, Vinod
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
CURRENT FILING DATE: 2002-03-08
CURRENT FILING DATE: 2002-03-08
PRIOR APPLICATION NUMBER: US 09/779,739
PRIOR FILING DATE: 2000-12-04
PRIOR APPLICATION NUMBER: US 09/653,450
PRIOR FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: US 09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US 09/598,042
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: US 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 63

```



```

RESULT 8
US-09-895-793-879
/ Sequence 879, Application US/09895793
/ Publication No. US20020192763A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Dillon, Devin C.
/ APPLICANT: Mitchell, Jennifer L.
/ APPLICANT: Harlocke, Susan L.
/ APPLICANT: Kato, Michael D.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Derrick
/ APPLICANT: Li, Samuel X.
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skelvy, Yaeli A.W.
/ APPLICANT: Hepler, William T.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Houghton, Raymond L.
/ APPLICANT: Foy, Teresa
/ APPLICANT: Fanger, Gary R.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.534C2
/ CURRENT APPLICATION NUMBER: US/09/895.793
/ PRIORITY FILING DATE: 2001-06-29
/ NUMBER OF SEQ ID NOS: 982
/ SOFTWARE: ParseSeq For Windows Version 3.0
/ SEQ ID NO 879
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-895-793-879
Query Match          100.0%; Score 116; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 1,7e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 REVHPLATSHOOFFKPIPLV 22
DB      92 REVHPLATSHOOFFKPIPLV 113

RESULT 9
US-09-895-814-879
/ Sequence 879, Application US/09895814
/ Publication No. US20020193296A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Mitchell, Jennifer L.
/ APPLICANT: Harlocke, Susan L.
/ APPLICANT: Kato, Michael D.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Carter, Derrick
/ APPLICANT: Li, Samuel X.
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skelvy, Yaeli A.W.
/ APPLICANT: Hepler, William T.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Houghton, Raymond L.

```

```

/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Houghton, Raymond L.
/ APPLICANT: Foy, Teresa
/ APPLICANT: Fanger, Gary R.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.427C26
/ CURRENT APPLICATION NUMBER: US/09/895.814
/ PRIORITY FILING DATE: 2001-06-29
/ NUMBER OF SEQ ID NOS: 330
/ SOFTWARE: ParseSeq For Windows Version 3.0
/ SEQ ID NO 879
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-895-814-879
Query Match          100.0%; Score 116; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 1,7e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 REVHPLATSHOOFFKPIPLV 22
DB      92 REVHPLATSHOOFFKPIPLV 113

RESULT 10
US-10-011-095-2
/ Sequence 2, Application US/10011095
/ Publication No. US20030045682A1
/ GENERAL INFORMATION:
/ APPLICANT: Alar, Daniel
/ APPLICANT: Leong, Kahm S.
/ APPLICANT: Saffran, Arthur B.
/ APPLICANT: Mitchell, Steve Chappell
/ TITLE OF INVENTION: ANTIBODIES IMMUNOSPECIFIC FOR STEAP1 (AS AMENDED)
/ FILE REFERENCE: 511562001610
/ CURRENT APPLICATION NUMBER: US/10/011.095
/ PRIORITY FILING DATE: 2001-12-06
/ NUMBER OF SEQ ID NOS: 1206
/ SOFTWARE: ParseSeq For Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ OTHER INFORMATION: DNA
US-10-011-095-2
Query Match          100.0%; Score 116; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 1,7e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 REVHPLATSHOOFFKPIPLV 22
DB      92 REVHPLATSHOOFFKPIPLV 113

RESULT 11
US-10-010-667a-2
/ Sequence 2, Application US/10010667a
/ Publication No. US20030055217A1
/ GENERAL INFORMATION:
/ APPLICANT: Alar, Daniel

```


QY 1 REVYHPLATSHQOYFYKIPILV 22
 Db 92 REVYHPLATSHQOYFYKIPILV 113

RESULT 15

US-09-802-520-1
 ; Sequence 1, Application US/09802520
 ; Publication No. US20020187472A1
 ; GENERAL INFORMATION:
 ; INVENTOR: Chen, Huel-Mei
 ; APPLICANT: Parle, Mary
 ; APPLICANT: Chen, Huel-Mei
 ; TITLE OF INVENTION: STEAP-RELATED PROTEIN
 ; FILE REFERENCE: PC-0037 US
 ; CURRENT APPLICATION NUMBER: US/09/802,520
 ; CURRENT FILING DATE: 2001-03-09
 ; NUMBER OF SEQ. IDS: 1
 ; SOFTWARE: PERL Program
 ; SEQ ID NO. 1
 ; LENGTH: 490
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURES:
 ; NAME/DEF: misc feature
 ; OTHER INFORMATION: Itcycle ID NO. US20020187472A1 7492446CD1
 US-09-802-520-1

Query Match 58.64; Score 68; DB 9; Length 490;
 Best Local Similarity 65.04; Pred. NO. 0.0083;
 Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 1 REVYHPLATSHQOYFYKIPILV 20
 Db 233 RQVHPLATSHQOYFYKIPILV 252

Search completed: March 26, 2003, 17:04:15
 Job time : 15.8372 secs

Genome version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 26, 2003, 16:44:40 / Search time 14.0698 Seconds
(Without alignments)
46.007 Million cell updates/sec

Title: US-10-010-667a-19

Perfect score: 116

Sequence: 1 REVHLPLATSHQOFKIPILV 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 15 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/aa/5A.COMB.pep:*
4: /cgn2_6/prodata/1/aa/5B.COMB.pep:*
5: /cgn2_6/prodata/1/aa/5A.COMB.pep:*
6: /cgn2_6/prodata/1/aa/5B.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	22	4 US-09-323-873a-19	Sequence 19, App1
2	116	100.0	339	4 US-09-323-873a-2	Sequence 2858, App1
3	46	39.7	195	4 US-09-134-001C-3548	Sequence 3548, App1
4	37.9	93.3	4	US-08-682-847-2	Sequence 2, App1
5	36.2	319	4	US-09-134-001C-3888	Sequence 1888, App1
6	36.2	1140	4	US-08-471-112a-3	Sequence 4, App1
7	42	36.2	1809	3 US-09-012-515a-12	Sequence 12, App1
8	42	36.2	1809	4 US-08-360-144a-12	Sequence 12, App1
9	42	36.2	1809	4 US-09-012-504a-12	Sequence 12, App1
10	42	36.2	1809	4 US-08-471-112a-3	Sequence 3, App1
11	42	36.2	2549	4 PCT-US95-06422-12	Sequence 12, App1
12	41.5	37.8	625	4 US-08-759-436-5	Sequence 5, App1
13	41.5	35.8	190	1 US-08-106-981-2	Sequence 2, App1
14	41.5	35.8	190	1 US-08-106-981-2	Sequence 2, App1
15	41.5	35.8	190	1 US-08-106-981-2	Sequence 2, App1
16	41.5	35.8	190	1 US-08-106-981-2	Sequence 2, App1
17	41	35.3	392	2 US-08-282-197C-60	Sequence 60, App1
18	41	35.3	392	2 US-08-282-197C-60	Sequence 60, App1
19	41	35.3	592	4 US-08-461-474-2	Sequence 2, App1
20	41	35.3	592	4 US-08-366-480-2	Sequence 2, App1
21	41	35.3	592	4 US-08-366-480-2	Sequence 2, App1
22	41	35.3	879	1 US-08-220-151-3	Sequence 3, App1
23	41	35.3	879	1 US-08-413-118-2	Sequence 2, App1
24	41	35.3	879	1 US-08-413-118-2	Sequence 2, App1
25	41	35.3	879	1 US-08-413-118-2	Sequence 2, App1
26	41	35.3	879	1 US-08-413-118-2	Sequence 2, App1
27	41	35.3	879	1 US-08-413-118-2	Sequence 2, App1

28	41	35.3	879	3 US-08-473-446-106	Sequence 106, App
29	40	34.5	168	4 US-09-134-001C-2858	Sequence 2858, App
30	40	34.5	168	4 US-09-134-001C-2858	Sequence 2858, App
31	40	34.5	206	4 US-08-213-787-2	Sequence 2, App1
32	40	34.5	303	1 US-08-185-432-5	Sequence 5, App1
33	40	34.5	353	4 US-09-461-474-4	Sequence 4, App1
34	40	34.5	485	4 US-09-291-023a-2	Sequence 2, App1
35	40	34.5	485	4 US-09-291-023a-2	Sequence 2, App1
36	40	34.5	485	4 US-09-291-023a-2	Sequence 2, App1
37	40	34.5	485	4 US-09-291-023a-2	Sequence 2, App1
38	40	34.5	485	4 US-09-291-023a-2	Sequence 2, App1
39	40	34.5	737	1 US-08-145-432-2	Sequence 2, App1
40	40	34.5	737	1 US-08-145-432-2	Sequence 2, App1
41	40	34.5	868	1 US-08-220-151-7	Sequence 4, App1
42	40	34.5	868	1 US-08-413-118-7	Sequence 7, App1
43	40	34.5	868	1 US-08-413-118-7	Sequence 7, App1
44	40	34.5	868	1 US-08-804-439a-21	Sequence 21, App1
45	40	34.5	868	3 US-08-473-446-7	Sequence 7, App1
46	40	34.5	868	3 US-08-473-446-7	Sequence 7, App1
47	40	34.5	868	3 US-08-473-446-7	Sequence 7, App1
48	40	34.5	868	3 US-08-473-446-7	Sequence 7, App1
49	40	34.5	868	3 US-08-473-446-7	Sequence 7, App1
50	40	34.5	868	3 US-08-473-446-7	Sequence 7, App1

ALIGNMENTS

```

RESULT 1
US-09-323-873a-19
/ Sequence 19, Application US/09323873a
/ General Information:
/ APPLICANT: Daniel E. Afar
/ APPLICANT: Rene S. Hubert
/ APPLICANT: Arthur B. Raitano
/ APPLICANT: Douglas C. Saffran
/ APPLICANT: Steve Chappell Mitchell
/ TITLE OF INVENTION: HUMAN TRANSMEMBRANE ANTIGENS
/ FILE REFERENCES: 129.16US02
/ CURRENT APPLICATION NUMBER: US/09/323,873A
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/093,183
/ PCT NO: 95/06-50
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 19
/ LENGTH: 22
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURES:
/ INFORMATION: STED-1-PEPTIDE
US-09-323-873a-19
Query Match 100.0%; Score 116; DB 4; Length 22;
Beat Local Similarity 100.0%; Pred. No. 4,3e+12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 REVHLPLATSHQOFKIPILV 22
1 REVHLPLATSHQOFKIPILV 22
RESULT 2
US-09-323-873a-2
/ Sequence 2858, Application US/09323873a
/ General Information:
/ APPLICANT: Daniel E. Afar
/ APPLICANT: Rene S. Hubert
/ APPLICANT: Arthur B. Raitano
/ APPLICANT: Douglas C. Saffran
/ APPLICANT: Steve Chappell Mitchell

```

```

/ TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
/ FILE REFERENCE: 128.160512
/ CURRENT FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/087,520
/ PRIOR FILING DATE: 1998-06-01
/ PRIOR APPLICATION NUMBER: 60/091,183
/ PRIOR FILING DATE: 1998-06-30
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: FASTEST for Windows Version 4.0
/ SEQ ID NOS: 1-32
/ LENGTH: 339
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-09-323-873A-2

Query Match      100.0%; Score 116; DB 4; Length 339;
Best Local Similarity 100.0%; Pred. No. 9,1e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 REVHPLATSHQGFYKIPILV 22
Db 92 REVHPLATSHQGFYKIPILV 113

RESULT 3
US-09-134-001C-3548
/ Sequence 3548, Application US/09134001C
/ Patent No. 6380370
/ GENERAL INFORMATION:
/ APPLICANT: Lynn Doucette-Stamm et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
/ TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
/ CURRENT APPLICATION NUMBER: US/09/134.001C
/ CURRENT FILING DATE: 1998-08-13
/ PRIOR APPLICATION NUMBER: US 60/064,964
/ PRIOR FILING DATE: 1997-11-08
/ PRIOR APPLICATION NUMBER: US 60/055,779
/ PRIOR FILING DATE: 1997-08-14
/ NUMBER OF SEQ ID NOS: 5674
/ SEQ ID NOS: 1-5674
/ LENGTH: 195
/ TYPE: PRT
/ ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3548

Query Match      39.7%; Score 46; DB 4; Length 195;
Best Local Similarity 44.4%; Pred. No. 6;
Matches 0; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Cy 1 REVHPLATSHQGFYK 18
Db 128 RSLSLPLTSLYQEWQOI 145

RESULT 4
US-08-682-847-2
/ Sequence 2, Application US/08682847
/ Patent No. 5658989
/ GENERAL INFORMATION:
/ APPLICANT: BABIUK, LORNE
/ APPLICANT: VAN DEN HURK, SYLVIA
/ APPLICANT: KAM, TIM
/ APPLICANT: FLETCHER, DAVID
/ TITLE OF INVENTION: RECOMBINANT BOVINE HERPESVIRUS TYPE 1
/ NUMBER OF SEQUENCES: 8
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: MORRISON & FORSTER
/ STREET: 755 PAGE MILL ROAD
/ CITY: PALO ALTO

```

```

/ STATE: CA
/ COUNTRY: USA
/ SEQ ID NOS: 1-9
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/682,847
/ FILING DATE: 1998-06-01
/ CLASSIFICATION: 2-196
/ ATTORNEY/AGENT INFORMATION:
/ NAME: PARK, FREDIE K.
/ REGISTRATION NUMBER: 35,636
/ REFERENCE/DOCKET NUMBER: 29310-20005.10
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 812-5600
/ TELEFAX: (415) 494-0792
/ TELEX: 706141
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 933 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-682-847-2

Query Match      37.9%; Score 44; DB 2; Length 933;
Best Local Similarity 40.0%; Pred. No. 71;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Cy 1 REVHPLATSHQGF 15
Db 660 REVERECTAHNRHVF 674

RESULT 5
US-09-357-251-24
/ Sequence 24, Application US/09357251
/ Patent No. 6271441
/ GENERAL INFORMATION:
/ APPLICANT: Palco, S. Carl
/ APPLICANT: Ramodu, Layo O.
/ APPLICANT: Ocozco, Buddy
/ APPLICANT: Schaeber, James S.
/ TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase
/ FILE REFERENCE: BB-1193
/ CURRENT APPLICATION NUMBER: US/09/357,251
/ CURRENT FILING DATE: 1998-07-30
/ EARLIER APPLICATION NUMBER: 60/093,530
/ EARLIER FILING DATE: July 21, 1998
/ NUMBER OF SEQ ID NOS: 37
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO: 24
/ LENGTH: 284
/ TYPE: PRT
/ ORGANISM: Trifolium aestivum
US-09-357-251-24

Query Match      36.2%; Score 42; DB 4; Length 284;
Best Local Similarity 46.7%; Pred. No. 39;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Cy 5 HPLATSHQGFYK 19
Db 55 HRARSHDTFLKAP 69

RESULT 6
US-09-134-001C-3888
/ Sequence 3888, Application US/09134001C
/ Patent No. 6380370

```



```

/ GENERAL INFORMATION:
/ APPLICANT: Lynn Doucette-Stamm et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
/ FILE REFERENCE: CITE-007
/ CURRENT APPLICATION NUMBER: US/09/134,001C
/ PRIOR FILING DATE: 1998-08-13
/ PRIOR APPLICATION NUMBER: US 60/064,964
/ PRIOR FILING DATE: 1997-11-08
/ PRIOR APPLICATION NUMBER: US 60/055,779
/ NUMBER OF SEQ ID NOS: 1997-08-14
/ SEQ ID NO 388
/ LENGTH: 319
/ TYPE: PRT
/ ORGANISM: Staphylococcus epidermidis
/ US-09-134-001C-388

Query Match          36.2% Score 42, DB 4, Length 319,
Best Local Similarity 36.1%, Pred. No. 45
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY      4  IHPLATSHQGYFYKPI 20
Db      249  IHKCYSDHQRFYVCSI 265

RESULT 7
US-08-471-112A-4
/ Sequence 4, Application US/08471112A
/ Patent No. 6313264
/ GENERAL INFORMATION:
/ APPLICANT: Molnar-Kimber, Katherine L.
/ APPLICANT: Fajali, Amedeo F.
/ APPLICANT: Peggiano, Thomas J.
/ APPLICANT: Natsopoulos, Ioannis J.
/ APPLICANT: Chen, Yangu
/ TITLE OF INVENTION: EFFECTOR PROTEINS OF RAPAMYCIN
/ NUMBER OF SEQUENCES: 23
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
/ Lerner, LLP
/ CITY: Washington
/ STATE: DC
/ COUNTRY: USA
/ ZIP: 20005-3315
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/471,112A
/ FILING DATE: 06-JUN-1995
/ CLASSIFICATION: 536
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/384,524
/ FILING DATE: 08-NOV-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/312,023
/ FILING DATE: 26-SEP-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/207,975
/ FILING DATE: 08-MAR-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Vincent, Matthew P.
/ REGISTRATION NUMBER: 26,709
/ TELEPHONE: 617-812-1000
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-408-4000
/ TELEFAX: 202-408-4400
/ INFORMATION FOR SEQ ID NO: 4:

```

```

/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1140 amino acids
/ STRANDS: 1
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-08-471-112A-4

Query Match          36.2% Score 42, DB 4, Length 1140,
Best Local Similarity 50.0%, Pred. No. 1,8e+02,
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      1  REVHPLATSHQGYFY 16
Db      274  RQDHPHLPFYHPOVY 289

RESULT 8
US-09-012-515A-12
/ Sequence 12, Application US/09012515A
/ Patent No. 6127521
/ GENERAL INFORMATION:
/ APPLICANT: Berlin, Vivian
/ APPLICANT: Chiu, Maria Isabel
/ APPLICANT: Cottrill, Guillaume
/ APPLICANT: Damagnez, Veronique
/ TITLE OF INVENTION: MONOSUPPRESSANT TARGET PROTEINS
/ NUMBER OF SEQUENCES: 3
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Foley, Hoag & Eliot LLP
/ STREET: One Post Office Square
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02110
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/012,515A
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/360,144
/ FILING DATE: 20-DEC-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Vincent, Matthew P.
/ REGISTRATION NUMBER: 26,709
/ TELEPHONE: 617-812-1000
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-812-1000
/ TELEFAX: 617-812-7000
/ INFORMATION FOR SEQ ID NO: 12:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1809 amino acids
/ TYPE: amino acid
/ MOLECULE TYPE: protein
/ US-09-012-515A-12

Query Match          36.2% Score 42, DB 3, Length 1809,
Best Local Similarity 50.0%, Pred. No. 3.1e+02,
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      1  REVHPLATSHQGYFY 16
Db      943  RQDHPHLPFYHPOVY 958

RESULT 9
US-08-360-144A-12

```

```

/ Sequence 12, Application US/08360144A
/ Patent No. 6150137
/ GENERAL INFORMATION:
/ APPLICANT: Chilly, Vivian
/ APPLICANT: Chilly, Michael
/ APPLICANT: Cotelarel, Guillaume
/ APPLICANT: Damagnez, Veronique
/ TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
/ NUMBER OF SEQUENCES: 35
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: FOLEY, HONG & ELIOT LLP
/ STREET: 1000 Post Office Square
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02109-2170
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/360,144A
/ FILING DATE: 20-DEC-1994
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Attorney, Michael F. 109
/ REGISTRATION NUMBER: 34,709
/ REFERENCE/DOCKET NUMBER: APV-036.02
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-832-1000
/ TELEFAX: 617-832-7000
/ INFORMATION FOR SEQ ID NO: 12:
/ SEQUENCE CHARACTERISTICS:
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-360-144A-12

Query Match      36.24; Score 42; DB 4; Length 1809;
Best Local Similarity 50.04; Pred. No. 3.1e+02;
Matches      8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      1 REVHPLATSHOOQFY 16
Db      943 RQDHLPLTVHPQTV 958

RESULT 10
US-09-012-504A-12
/ Sequence 12, Application US/09012504A
/ Patent No. 6464974
/ GENERAL INFORMATION:
/ APPLICANT: Beilín, V.
/ APPLICANT: Chiu, I.
/ APPLICANT: Cotelarel, G.
/ APPLICANT: Damagnez, V.
/ TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
/ CURRENT APPLICATION NUMBER: US/09/012,504A
/ PRIOR FILING DATE: 1998-01-23
/ PRIOR APPLICATION NUMBER: 08/360,144
/ PRIOR APPLICATION DATE: 1994-12-20
/ NUMBER OF SEQ ID NOS: 35
/ SOFTWARE: Patent In Release 3.1
/ SEQ ID NO 12
/ LENGTH: 1809
/ TYPE: PRT
/ ORGANISM: Mammalian
/ US-09-012-504A-12

```

```

Query Match      36.24; Score 42; DB 4; Length 1809;
Best Local Similarity 50.04; Pred. No. 3.1e+02;
Matches      8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      1 REVHPLATSHOOQFY 16
Db      943 RQDHLPLTVHPQTV 958

RESULT 11
US-08-471-112A-3
/ Sequence 3, Application US/08471112A
/ Patent No. 6313264
/ GENERAL INFORMATION:
/ APPLICANT: Molnar-Kimber, Katherine L.
/ APPLICANT: Fallini, Amedeo F.
/ APPLICANT: Casagrande, Thomas J.
/ APPLICANT: Karamitros, Koji
/ APPLICANT: Chilly, Michael
/ TITLE OF INVENTION: EFFECTOR PROTEINS OF RABMYCIN
/ NUMBER OF SEQUENCES: 23
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
/ STREET: 1100 I Street, N.W.
/ CITY: Washington
/ STATE: DC
/ COUNTRY: USA
/ ZIP: 20005-3315
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/471,112A
/ FILING DATE: 06-JUN-1995
/ CLASSIFICATION: 536
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/384,524
/ FILING DATE: 08-MAR-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/312,023
/ FILING DATE: 26-SEP-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/207,975
/ FILING DATE: 08-MAR-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Finnegan, Henderson, Farabow, Garrett &
/ REGISTRATION NUMBER: 36,276
/ REFERENCE/DOCKET NUMBER: 01142.0058-00000
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-408-4400
/ TELEFAX: 202-408-4400
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ TYPE: amino acid
/ STRANDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-08-471-112A-3

Query Match      36.24; Score 42; DB 4; Length 2549;
Best Local Similarity 50.04; Pred. No. 4.5e+02;
Matches      8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      1 REVHPLATSHOOQFY 16
Db      1683 RQDHLPLTVHPQTV 1698

```

```

RESULT 12
PCT-US95-06722-12
Sequence 12, Application PCT/US9506722
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Immunosuppressant Target Proteins
NUMBER OF SEQUENCES: 25
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
FILING DATE: 20-DEC-1994
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06722
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/250,795
FILING DATE: 27-MAY-1994
PREVIOUS APPLICATION DATA:
APPLICATION NUMBER: US 08/250,795
FILING DATE: 20-DEC-1994
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2549 amino acids
TYPE: amino acid
TOPODOGY: linear
NO. OF SEQUENCES: 1
PCT-US95-06722-12
Query Match 36.24; Score 42; DB 5; Length 2549;
Best Local Similarity 50.04; Pred. No. 4, Seq 02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
Qy 1 REVPLATSHQOYF 16
Db 1683 RQDHPPLFVHPQVY 1698

```

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 616-5600
FACSIMILE: (312) 616-7000
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 625 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-581-148C-18
Query Match 35.84; Score 41.5; DB 3; Length 625;
Best Local Similarity 52.44; Pred. No. 1, Seq 02;
Matches 11; Conservative 2; Mismatches 5; Indels 3; Gaps 2;
Qy 3 VHPPLATSHQOYF--YKIPIL 21
Db 174 VHPPLA-SHIAVFLPALPL 193
RESULT 14
US-08-759-436-3
Sequence 3, Application US/08759436
Patent No. 6437218
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: DNA fragment encoding a protein
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/759,436
FILING DATE: 5-Dec-1996
ATTORNEY/AGENT INFORMATION:
NAME: Cord, Janet I.
REGISTRATION NUMBER: 33,778
TELEPHONE: 212-246-8959
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-246-8959
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 625 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-759-436-3
Query Match 35.84; Score 41.5; DB 4; Length 625;
Best Local Similarity 52.44; Pred. No. 1, Seq 02;
Matches 11; Conservative 2; Mismatches 5; Indels 3; Gaps 2;
Qy 3 VHPPLATSHQOYF--YKIPIL 21
Db 174 VHPPLA-SHIAVFLPALPL 193
RESULT 15
US-08-759-436-5

```

Search completed: March 26, 2003, 16:51:39
 CPU time : 16.0696 secs

Sequence 5, Application US/08759436
 Patent No. 6437218
 GENERAL INFORMATION:
 APPLICATION:
 TITLE OF INVENTION: DNA fragment encoding a protein
 TITLE OF INVENTION: Involved in
 TITLE OF INVENTION: fatty aldehyde decarboxylase activity, recombinant
 TITLE OF INVENTION: molecules comprising said fragment and a method for
 TITLE OF INVENTION: obtaining transformed bacterial cells and plants
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESS: 26 West 61st Street
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10023
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/759,436
 FILING DATE: 5-Dec-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: COLO, James L.
 ADDRESS: 11177
 TELEPHONE: 212-708-1935
 TELEFAX: 212-246-8959
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 453 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Arabidopsis thaliana
 DEVELOPMENTAL STAGE: light and dark grown seedlings;
 IMMEDIATE SOURCE: rosette and whole plants
 LIBRARY: PRL2
 FEATURE:
 NAME/KEY: Protein
 LOCATION: 1-453
 PUBLICATION INFORMATION:
 AUTHORS: Artter, Mark G.M.
 AUTHORS: Keijzer, Christian J.
 AUTHORS: Stekema, Willem J.
 TITLE: Molecular characterization of the CER1 gene
 TITLE: of Arabidopsis involved in epicuticular wax
 TITLE: biosynthesis and pollen fertility
 JOURNAL: Plant Cell
 VOLUME: 7
 ISSUE: 12
 PAGES: 2115-2127
 DATE: december.1995
 RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 625
 US-08-759-436-5

Query Match 35.84; Score 41.5; DB 4; Length 625;
 Best Local Similarity 52.44; Prnd No. 1.1e+02;
 Matches 11; Conservative 2; Mismatches 5; Indels 3; Gaps 2;
 QY 3 VHPPLATSHOQYF--VHPIL 21
 DB 174 VHPPLA-SHIAIFLIPALILL 193